

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 15:33:46 ; Search time 3409 Seconds
(without alignments)
10908.866 Million cell updates/sec

Title: US-09-889-733b-1

Perfect score: 858
Sequence: 1 atggcggcctcgccaccc.....acaagaaggtagactcctga 858

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hcg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_atgo_hum.*
- 40: em_atgo_mus.*
- 41: em_atgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	858	100.0	858	6	BD251986	BD251986 Screen me
2	858	100.0	858	6	AX032835	AX032835 Sequence
3	851.6	99.3	1624	9	BC033680	BC033680 Homo sapi
4	851.6	99.3	1633	6	BD094073	BD094073 Shear str
5	851.6	99.3	1633	6	AB001915	AB001915 Homo sapi
6	851.6	99.3	4208	6	AX548046	AX548046 Sequence
7	739.4	86.2	3737	9	BC043235	BC043235 Homo sapi
8	717.2	83.6	1120	10	BC034505	BC034505 Mus muscu
9	715.6	83.4	3008	6	AX683109	AX683109 Sequence
10	715.6	83.4	3008	10	D86041	D86041 Rattus norv
11	579.6	67.6	1150	5	AB095027	AB095027 Gallus ga
12	555.4	64.7	3908	9	HS080822	HS080822 Homo sapi
13	462.4	53.9	2289	10	AX128909	AX128909 Mus muscu
14	457.8	53.4	1344	5	BC056074	BC056074 Xenopus l
15	381.8	44.5	3420	10	BC060266	BC060266 Mus muscu
16	328.8	38.3	2895	9	AK098405	AK098405 Homo sapi
17	300.2	35.0	104071	2	AL162260	AL162260 Homo sapi
18	300.2	35.0	158405	2	AC019261	AC019261 Homo sapi
19	300.2	35.0	168953	9	AL360219	AL360219 Human DNA
20	260.2	30.3	186434	2	AC126412	AC126412 Mus muscu
21	258.6	30.1	228940	2	AC126890	AC126890 Rattus no
22	216.2	25.2	1263	9	BC001435	BC001435 Homo sapi
23	216.2	25.2	1342	9	AK098671	AK098671 Homo sapi
24	216.2	25.2	1376	6	AX011712	AX011712 Sequence
25	216.2	25.2	1376	6	BD226323	BD226323 Pancreat
26	216.2	25.2	1709	9	AK026191	AK026191 Homo sapi
27	214.6	25.0	858	6	BD251987	BD251987 Screen me
28	214.6	25.0	858	6	AX032837	AX032837 Sequence
29	214.6	25.0	1351	6	AX779950	AX779950 Sequence
30	214.6	25.0	1351	9	AF087894	AF087894 Homo sapi
31	202.8	23.6	1310	10	BC003328	BC003328 Mus muscu
32	193.2	22.5	1270	10	AF004106	AF004106 Mus muscu
33	189.8	22.1	1228	6	AX780306	AX780306 Sequence
34	185.6	21.6	1397	6	AX79492	AX79492 Sequence
35	148	17.2	83946	9	HS123M24	HS123M24 Human DNA
36	122.4	14.3	58539	2	AC137903	AC137903 Mus muscu
37	122.4	14.3	200007	2	AC123684	AC123684 Mus muscu
38	119.8	14.0	21211	2	AC094482	AC094482 Rattus no
39	107.4	12.5	1377	3	AY060959	AY060959 Drosophil
40	97.8	11.4	186434	2	AC126412	AC126412 Mus muscu
41	93.2	10.9	21211	2	AC094482	AC094482 Rattus no
42	92.8	10.8	110000	2	AC144263	Continuation (3 of
43	92.8	10.8	304208	2	AC143028	AC143028 Macaca mu
44	85.6	10.0	626	11	BV064692	BV064692 S212P6014
45	85	9.9	420	6	AX898477	AX898477 Sequence

ALIGNMENTS

RESULT 1
BD251986
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

BD251986
Screen method.
BD251986
BD251986.1 GI:33061756
JP 2002535001-A/1.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 858)
Vallance,P.J.T., Leiper,J.M., Whitley,G.S.J. and Charles,I.G.
Screen method
Patent: JP 2002535001-A 1 22-OCT-2002;

858 bp DNA linear PAT 17-JUL-2003

Pred. No. is the number of results predicted by chance to have a

UNIVERSITY COLLEGE LONDON
OS Homo sapiens (human)
EN JP 2002535001-A/1
FD 22-OCT-2002
PF 26-JAN-2000 JP 2000596131
PR 26-JAN-1999 GB 9901705.5, 04-JUN-1999 GB 9913066.8 PI
PATRICK JOHN THOMPSON VALLANCE, JAMES MITCHELL LEIPER, GUY ST PI
JOHN WHITLEY,
PI IAN GEORGE CHARLES
PC C12N15/09, A01K67/027, A61K31/198, A61K38/00, A61K45/00, A61P3/06,
PC A61P9/02,
PC A61P9/04, A61P9/10, A61P9/12, A61P13/12, A61P25/06, A61P25/18, PC
A61P25/28,
PC A61P29/00, A61P31/04, A61P35/00, C07K16/40, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12N9/78, C12Q1/02, C12Q1/34, G01N33/15, G01N33/50 PC
C12N15/00, C12N5/00,
PC A61K37/02
CC Screen method
FH Key
FT CDS Location/Qualifiers
1. 858
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
Query Match 100.0%; Score 858; DB 6; Length 858;
Best Local Similarity 100.0%; Pred. No. 8.1e-132; Indels 0; Gaps 0;
Matches 858; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
1 ATGCGCGGCTCGGCACCCCTCGGCTTCGGCGGCGGCACCCACCCCTGTGTGGCGGCG 60
2 ATGCGCGGCTCGGCACCCCTCGGCTTCGGCGGCGGCACCCACCCCTGTGTGGCGGCG 60
3 CTACCCGAGTCGCTGCGCAGCAGCGCTGAGAGCGCCAGCGCGGAGGAGGTGACGTC 120
4 CTACCCGAGTCGCTGCGCAGCAGCGCTGAGAGCGCCAGCGCGGAGGAGGTGACGTC 120
5 GCCCGCGGGAACGGCAGCAGCAGCTACGTGGGCGTCTGGCGAGCAAGCTGGGGCTG 180
6 GCCCGCGGGAACGGCAGCAGCAGCTACGTGGGCGTCTGGCGAGCAAGCTGGGGCTG 180
7 CAGGTGTGAGTGTGGCGGCGCAGAGCCTTCGGACTCGCTTCGTGGAGGACGTC 240
8 CAGGTGTGAGTGTGGCGGCGCAGAGCCTTCGGACTCGCTTCGTGGAGGACGTC 240
9 GCCGTGTGCGGAGGAGCGCCCTCATCCCGACCGCGGCGCGCGCGGCGGAGGAG 300
10 GCCGTGTGCGGAGGAGCGCCCTCATCCCGACCGCGGCGCGCGCGGCGGAGGAG 300
11 GAGTTGACATGATGAAAGAGCATTAGAAAACCTTCAGTCAATATAGTAGAGTAAA 360
12 GAGTTGACATGATGAAAGAGCATTAGAAAACCTTCAGTCAATATAGTAGAGTAAA 360
13 GATGAAATCCAACTTTAGATGCGGAGATGTTTATTCACAGGCGAGAGATTTTGTG 420
14 GATGAAATCCAACTTTAGATGCGGAGATGTTTATTCACAGGCGAGAGATTTTGTG 420
15 GGCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGCTGATCTTTAAGGAC 480
16 GGCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGCTGATCTTTAAGGAC 480
17 TATGCACTCCACAGTCCAGTGGGAGTGGGTTCGATTTGAAGATTTCTGCAGCATG 540
18 TATGCACTCCACAGTCCAGTGGGAGTGGGTTCGATTTGAAGATTTCTGCAGCATG 540
19 GCTGGGCTTAACCTGATCGCAATTTGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATC 600
20 GCTGGGCTTAACCTGATCGCAATTTGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATC 600
21 ATGCACAGATGAGTGTACCCAGCCGCTACGACAACTCACTGTGCTCATGACATAGCAGCA 660

Db 601 ATGCACAGATGAGTGTACCCAGCCGCTACGCAAACTCACTGTGCTGATGACATAGCAGCA 660
Qy 661 AACTGTATATATTAATATATCCCAAAAGGGCAGCTTGTGTCGACCCGAAACCCCGGAA 720
Db 661 AACTGTATATATTAATATATCCCAAAAGGGCAGCTTGTGTCGACCCGAAACCCCGGAA 720
Qy 721 GAGTATCCAGAAAGTCAAAAGGTTTATGAGAACTGAAGACCATATGCTGATCCCCGCTG 780
Db 721 GAGTATCCAGAAAGTCAAAAGGTTTATGAGAACTGAAGACCATATGCTGATCCCCGCTG 780
Qy 781 AGCATGCTGAACTGAAAGAGGTGATGGCTGCTACCTGCTGCTCAGTTTAAITTAAC 840
Db 781 AGCATGCTGAACTGAAAGAGGTGATGGCTGCTACCTGCTGCTCAGTTTAAITTAAC 840
Qy 841 AAGAAGGTAGACTCCTCA 858
Db 841 AAGAAGGTAGACTCCTCA 858

RESULT 2
AX032835
LOCUS AX032835
DEFINITION Sequence 1 from Patent WO004488.
ACCESSION AX032835
VERSION AX032835.1 GI:10279809
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Leiper, J.M., Whitley, G.S., Charles, I.G. and Vallance, P.J.
Screen method
Patent: WO 004488-A 1 03-AUG-2000;
PATRICK JAMES MITCHELL (GB); WHITLEY GUY ST JOHN (GB); UNIV LONDON
(GB); CHARLES IAN GEORGE (GB); VALLANCE PATRICK JOHN THOMPSON
(GB)

FEATURES
source
1. 858
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1. 858
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC09947.1"
/db_xref="GI:10279810"
/translation="MAGLHPSAFGRATHAVVRALPESLCOHALRSKAGBEVDVARAE
RQHLVYVGLSGKLQVVELPADSLPDCVFEDVAVVCEETALITRPAPSRKREV
DMKEALEKLQNLIVEMKEDNATLDGDLVLTGREFVGLSKRTNORGAAILADTPD
YAVSTVPADGLHLKSPCSMAGNLIATGSESNOKALKIMQMSDHRVYKLTVPDI
AANCYINIPNKGHVLLHRTPEEPYPSAKVTEKLKDHMLIPVSNSELEKYDGLLTCCS
VLNKKVDS"

ORIGIN
Query Match 100.0%; Score 858; DB 6; Length 858;
Best Local Similarity 100.0%; Pred. No. 8.1e-132; Indels 0; Gaps 0;
Matches 858; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCGCGGCTCGGCACCCCTCGGCTTCGGCGGCGGCACCCACCCCTGTGTGGCGGCG 60
Db 1 ATGCGCGGCTCGGCACCCCTCGGCTTCGGCGGCGGCACCCACCCCTGTGTGGCGGCG 60
Qy 61 CTACCCGAGTCGCTGCGCAGCAGCGCTGAGAGCGCCAGCGCGGAGGAGGTGACGTC 120
Db 61 CTACCCGAGTCGCTGCGCAGCAGCGCTGAGAGCGCCAGCGCGGAGGAGGTGACGTC 120
Qy 121 GCCCGCGGGAACGGCAGCAGCAGCTACGTGGGCGTCTGGCGAGCAAGCTGGGGCTG 180
Db 121 GCCCGCGGGAACGGCAGCAGCAGCTACGTGGGCGTCTGGCGAGCAAGCTGGGGCTG 180

family contains glycine (EC:2.1.4.1) and inosamine (EC:2.1.4.2) amidinotransferases, enzymes involved in creatine and streptomycin biosynthesis respectively. This family also includes arginine deiminases, EC:3.5.3.6. These enzymes catalyse the reaction: arginine + H2O <=> citrulline + NH3. Also found in this family is the Streptococcus anti tumor glycoprotein
/db_xref="CDD:pfam02274"

ORIGIN		Query Match	99.3%; Score 851.6; DB 9; Length 1624;	
		Best Local Similarity	99.5%; Pred. No. 9.1e-131;	
		Mismatches	854; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
Qy	1	ATGGCGCGCTCGGCACCCCTTCGCGCTTGGCCCGGCGACCCACCGCGTGGTGGCGGC	60	
Db	100	ATGGCGCGCTCGGCACCCCGCCCTTGGCCCGGCGACCCACCGCGTGGTGGCGGC	159	
Qy	61	CTACCGGAGTCTCTGCGCAGCAGCGCTGAGAAGCGCAAGGGGAGAGGTGACGTC	120	
Db	160	CTACCGGAGTCTCTGCGCAGCAGCGCTGAGAAGCGCAAGGGGAGAGGTGACGTC	219	
Qy	121	GCCCGCGGGAACGCGACAGCTCTACGTTGGCGCTGCTGGCGAGCAAGCTGGGCTG	180	
Db	220	GCCCGCGGGAACGCGACAGCTCTACGTTGGCGCTGCTGGCGAGCAAGCTGGGCTG	279	
Qy	181	CAGGTGCTGAGCTCCGCGCGACAGAGCCTTCGGGACTCGTCTTCTGCGAGGAGCTG	240	
Db	280	CAGGTGCTGAGCTCCGCGCGACAGAGCCTTCGGGACTCGTCTTCTGCGAGGAGCTG	339	
Qy	241	GCGGTGCTGCGAGAGACGCGCCCTCATCACCGACCCGGGCGCCGAGCGGAGGAAG	300	
Db	340	GCGGTGCTGCGAGAGACGCGCCCTCATCACCGACCCGGGCGCCGAGCGGAGGAAG	399	
Qy	301	GAGTTGACATGATGAAGAAGCATTAGAAAACCTTCAGCTCAATATAGTAGAGTGA	360	
Db	400	GAGTTGACATGATGAAGAAGCATTAGAAAACCTTCAGCTCAATATAGTAGAGTGA	459	
Qy	361	GATGAATGCAACTTAGATGCGGAGATGTTTATTACAGGCGAGAGATTTTGTG	420	
Db	460	GATGAATGCAACTTAGATGCGGAGATGTTTATTACAGGCGAGAGATTTTGTG	519	
Qy	421	GGCTTTCCAAAGGACAAATCAAGAGGTGCTGAAATCTTGCTGTGATCTTTTAAGAC	480	
Db	520	GGCTTTCCAAAGGACAAATCAAGAGGTGCTGAAATCTTGCTGTGATCTTTTAAGAC	579	
Qy	481	TATGAGTCTCCACAGTGCAGTGCAGATGGTTCATTTGAGAGTTTCTGAGCATG	540	
Db	580	TATGAGTCTCCACAGTGCAGTGCAGATGGTTCATTTGAGAGTTTCTGAGCATG	639	
Qy	541	GCTGGGCTTAACCTGATCGCAATGGGTCTAGTGAATCTGCACAGAAGCCCTTAAGATC	600	
Db	640	GCTGGGCTTAACCTGATCGCAATGGGTCTAGTGAATCTGCACAGAAGCCCTTAAGATC	699	
Qy	601	ATGCAACAGATGAGTGACCCGCTACGCAAACTCAGTGTGCTGTGATGATAGCAGCA	660	
Db	700	ATGCAACAGATGAGTGACCCGCTACGCAAACTCAGTGTGCTGTGATGATAGCAGCA	759	
Qy	661	AACTGTATATCTAAATATCCCAACAAAGGGCACGCTTGTCTGCACCGAACCCCGAA	720	
Db	760	AACTGTATATCTAAATATCCCAACAAAGGGCACGCTTGTCTGCACCGAACCCCGAA	819	
Qy	721	GAGTATCCAGAAAGTGAAGGTTTATGAGAACTGAAGGACCATATGCTGATCCCGTG	780	
Db	820	GAGTATCCAGAAAGTGAAGGTTTATGAGAACTGAAGGACCATATGCTGATCCCGTG	879	
Qy	781	AGCATGCTGAACTGGAAGGTTGATGGGCTGCTCACTGCTGCTCAGTTTAAATTAAC	840	
Db	880	AGCATGCTGAACTGGAAGGTTGATGGGCTGCTCACTGCTGCTCAGTTTAAATTAAC	939	
Qy	841	AAGAAGGTAGATCTCTGA	858	
Db	940	AAGAAGGTAGATCTCTGA	957	

RESULT 4	BD094073	1633 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD094073	1633 bp	DNA	linear	PAT 27-AUG-2002
DEFINITION	Shear stress-responsive DNAs.				
ACCESSION	BD094073	GI:22639661			
VERSION	WO 0125427-A/34				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1633)				
AUTHORS	Nojima, H., Yoshisue, H., Obayashi, M., Ota, T., Kawabata, A., Sakurada, K., Kuga, T., Sekine, S., Nakamura, Y. and Sugan, S.				
TITLE	Shear stress-responsive DNAs				
JOURNAL	Patent: WO 0125427-A 34 12-APR-2001; KYOMA HAKO KOGYO CO LTD, HIROSHI NOJIMA, HAJIME YOSHISUE, MASAYA ODAYASHI, TOSHIO OTA, AYAKO KAWABATA, KAZUHIRO SAKURADA, TETSURO KUGA, SUSUMU SEKINE, YUSUKE NAKAMURA, SUMIO SUGANO				
COMMENT	OS Homo sapiens (human) FN WO 0125427-A/34 PD 12-APR-2001 PF 02-OCT-2000 WO 2000JP006840 PI 01-OCT-1999 JP 99P 280976 PI HIROSHI NOJIMA, HAJIME YOSHISUE, MASAYA ODAYASHI, TOSHIO OTA, PI AYAKO KAWABATA, TETSURO KUGA, SUSUMU SEKINE, YUSUKE NAKAMURA, PI KAZUHIRO SAKURADA, TETSURO KUGA, SUSUMU SEKINE, YUSUKE NAKAMURA, PI SUMIO SUGANO PC C12N15/12, C07K14/435, C07K16/18, C12P21/02, C12Q1/68, A61K38/00, PC A61K39/395, PC A61K48/00, A61P9/10, G01N33/50, G01N33/53 CC				
FT	Key	Location/Qualifiers			
CDS	Location/Qualifiers	(323)..(1177).			
FEATURES	Source	1..1633			
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Query Match	99.3%; Score 851.6; DB 6; Length 1633;				
Best Local Similarity	99.5%; Pred. No. 9.1e-131;				
Mismatches	854; Conservative 0; Mismatches 4; Indels 0; Gaps 0;				
Qy	1	ATGGCGCGCTCGGCACCCCTTCGCGCTTGGCCCGGCGACCCACCGCGTGGTGGCGGC	60		
Db	323	ATGGCGCGCTCGGCACCCCGCCCTTGGCCCGGCGACCCACCGCGTGGTGGCGGC	382		
Qy	61	CTACCGGAGTCTCTGCGCAGCAGCGCTGAGAAGCGCAAGGGGAGAGGTGACGTC	120		
Db	383	CTACCGGAGTCTCTGCGCAGCAGCGCTGAGAAGCGCAAGGGGAGAGGTGACGTC	442		
Qy	121	GCCCGCGGGAACGCGACAGCTCTACGTTGGCGCTGCTGGCGAGCAAGCTGGGCTG	180		
Db	443	GCCCGCGGGAACGCGACAGCTCTACGTTGGCGCTGCTGGCGAGCAAGCTGGGCTG	502		
Qy	181	CAGGTGCTGAGTTCGCGCGGCGACGAGCGCTTCGCGACTGCGTCTTCGTTGAGGACGTC	240		
Db	503	CAGGTGCTGAGTTCGCGCGGCGACGAGCGCTTCGCGACTGCGTCTTCGTTGAGGACGTC	562		
Qy	241	GCGTGGTGGTGGCGAGAGCGCCCTCATCACCGACCCGGGCGCGGAGCGGAGGAAG	300		
Db	563	GCGTGGTGGTGGCGAGAGCGCCCTCATCACCGACCCGGGCGCGGAGCGGAGGAAG	622		
Qy	301	GAGTTGACATGATGAAGAAGCATTAGAAAACTTCAGCTCAATATAGTAGAGTGA	360		
Db	623	GAGTTGACATGATGAAGAAGCATTAGAAAACTTCAGCTCAATATAGTAGAGTGA	682		
Qy	361	GATGAAAATGCAACTTAGATGCGGAGATGTTTATTACAGGCGAGAGATTTTGTG	420		

Db 683 GATGAAATGCAACTTTAGATGGCGAGATGTTTATTTCACAGGCAGAGAAATTTTGTG 742
QY 421 GGCCTTTCCAAAGGACAAATCAACAGAGTGTCTGAAATCTTGGCTGATACATTTTAAAGAC 480
Db 743 GGCCTTTCCAAAGGACAAATCAACAGAGTGTCTGAAATCTTGGCTGATACATTTTAAAGAC 802
QY 481 TATGAGTCTCCACAGTCCAGTGCAGATGGTGTGCAATTTTGAAGAGTTTCTGCAGCATG 540
Db 803 TATGAGTCTCCACAGTGCAGTGCAGATGGTGTGCAATTTTGAAGAGTTTCTGCAGCATG 862
QY 541 GCTGGGCTTACCTGATCGCAATTTGGTCTAGTGAATCTGCACAGAGGCGCTTAAAGATC 600
Db 863 GCTGGGCTTACCTGATCGCAATTTGGTCTAGTGAATCTGCACAGAGGCGCTTAAAGATC 922
QY 601 ATGCAACAGATGAGTGCAGCACCAGCTACGACAACTCACTGTGCTGTGATGACATAGCAGCA 660
Db 923 ATGCAACAGATGAGTGCAGCACCAGCTACGACAACTCACTGTGCTGTGATGACATAGCAGCA 982
QY 661 AACTGTATATATCTAAATATCCCAACAAAGGCGAGCTGTGTCACCGAACCCCGGAA 720
Db 983 AACTGTATATATCTAAATATCCCAACAAAGGCGAGCTGTGTCACCGAACCCCGGAA 1042
QY 721 GAGTATCCAGAAAGTGCAGAAAGTGTATGAGAACTGAAAGGACCATATGCTGATCCCGGTG 780
Db 1043 GAGTATCCAGAAAGTGCAGAAAGTGTATGAGAACTGAAAGGACCATATGCTGATCCCGGTG 1102
QY 781 AGCATGTCTGAATCTGAAAGTGTGATGGTGTCTGCTACCTGCTGCTAGTTTAAATTAAC 840
Db 1103 AGCATGTCTGAATCTGAAAGTGTGATGGTGTCTGCTACCTGCTGCTAGTTTAAATTAAC 1162
QY 841 AAGAAGGTAGACTCTCTGA 858
Db 1163 AAGAAGGTAGACTCTCTGA 1180

fish version history

RESULT 5
AB001915
LOCUS AB001915 1633 bp mRNA linear PRI 21-SEP-2000
DEFINITION Homo sapiens mRNA for NG,NG-dimethylarginine dimethylaminohydrolase, complete cds.
ACCESSION AB001915
VERSION AB001915.1 GI:4160665
KEYWORDS NG,NG-dimethylarginine dimethylaminohydrolase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
Kimoto,M., Miyatake,S., Sasagawa,T., Yamashita,H., Okita,M., Oka,T., Ogawa,T. and Tsuji,H.
Purification, cDNA cloning and expression of human NG,NG-dimethylarginine dimethylaminohydrolase
Eur.J. Biochem. 258 (2), 863-868 (1998)
95089653
PUBMED 9874257
REFERENCE 2 (bases 1 to 1633)
Kimoto,M.
Direct Submission
Submitted (15-MAR-1997) Masumi Kimoto, Okayama Prefectural University, Faculty of Health and Welfare Science, Department of Nutritional Science, Kuboki 111, Soja, Okayama 719-1197, Japan
(E-mail:kimoto@fhw.oka.-pu.ac.jp, Tel:0866-94-2158, Fax:0866-94-2158)
FEATURES
1. 1633
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="Kidney"
323..1180
/EC number="3.5.3.18"
/codon_start=1
/product="NG,NG-dimethylarginine dimethylaminohydrolase"

CDS

/protein_id="BAA37117.1"
/db_xref="GI:4160666"
/translation="MAGLGHBAFPAFPAHVVRAVPESLGOHALSAKGEVVDVAEAE
RQHLVGVGLGSKLGLQVVELPADSLPDCVFVEDVAVVCEBTALITRGASRRKEV
DMKRELEKLNIVEMKDEENATDGGDLVFTGREFVGLSKATNGRAEILADTPKD
YAVSTVPVADGLHLKFCFSCMAGNLIAGSSBSAQKALKIMQMSDHRVYKLVTPDPI
AANCYVINTPNKHVILLHRTPEYPESAKVYKELKDHMLIPVMSSELEKVDGLLTCCS
VLINKKVDs"
ORIGIN
Query Match 99.3%; Score 851.6; DB 9; Length 1633;
Best Local Similarity 99.8%; Pred. No. 9.1e-131;
Matches 854; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGCGCGGCTCGCGCCACCCCTCCGCTTCGCGCGGCGCCACCCACCGCTGTGTGCGGGCG 60
Db 323 ATGCGCGGCTCGCGCCACCCCTCCGCTTCGCGCGGCGCCACCCACCGCTGTGTGCGGGCG 382
QY 61 CTACCGAGTGTCTGCGCAGCAGCGCTGAGAAAGCGCCAAAGCGCAGAGGTGAGCGTC 120
Db 383 CTACCGAGTGTCTGCGCAGCAGCGCTGAGAAAGCGCCAAAGCGCAGAGGTGAGCGTC 442
QY 121 GCCCGCGGAAAGCGCAGCAGCAGCTTACGTGGGGGTGTGCGGCGAGCAAGCTGGGGCTG 180
Db 443 GCCCGCGGAAAGCGCAGCAGCAGCTTACGTGGGGGTGTGCGGCGAGCAAGCTGGGGCTG 502
QY 181 CAGGTGTGAGTGTGCGCGCGCAGAGAGCTTTCGCGAGCTGTCTTTCGTGAGAGCGTG 240
Db 503 CAGGTGTGAGTGTGCGCGCGCAGAGAGCTTTCGCGAGCTGTCTTTCGTGAGAGCGTG 562
QY 241 GCCGTGTGTGCGAGGAGCGGCGCTCATCACCAGCCGCGGCGCGCGAGCGGAGGAG 300
Db 563 GCCGTGTGTGCGAGGAGCGGCGCTCATCACCAGCCGCGGCGCGCGAGCGGAGGAG 622
QY 301 GAGGTTGACATGATGAAAGAGCATTTAGAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
Db 623 GAGGTTGACATGATGAAAGAGCATTTAGAAAACTTCAGCTCAATATAGTAGAGATGAAA 682
QY 361 GATGAAATGCAACTTTAGTGTGCGGAGATGTTTATTCACAGGCGAGAGATTTTGTG 420
Db 683 GATGAAATGCAACTTTAGTGTGCGGAGATGTTTATTCACAGGCGAGAGATTTTGTG 742
QY 421 GGCCTTTCCAAAGGACAAATCAACAGAGTGTCTGAAATCTTGGCTGATACATTTTAAAGAC 480
Db 743 GGCCTTTCCAAAGGACAAATCAACAGAGTGTCTGAAATCTTGGCTGATACATTTTAAAGAC 802
QY 481 TATGAGTCTCCACAGTGCAGTGGCAGATGGGTTCATTTGAAGAGTTTCTGCAGCATG 540
Db 803 TATGAGTCTCCACAGTGCAGTGGCAGATGGGTTCATTTGAAGAGTTTCTGCAGCATG 862
QY 541 GCTGGGCTTAACTGATCGCAATTTGGGTCTAGTGAATCTGCACAGAGGCGCTTAAAGATC 600
Db 863 GCTGGGCTTAACTGATCGCAATTTGGGTCTAGTGAATCTGCACAGAGGCGCTTAAAGATC 922
QY 601 ATGCAACAGATGAGTGCAGCACCAGCTACGACAACTCACTGTGCTGTGATGACATAGCAGCA 660
Db 923 ATGCAACAGATGAGTGCAGCACCAGCTACGACAACTCACTGTGCTGTGATGACATAGCAGCA 982
QY 661 AACTGTATATCTAAATATCCCAACAAAGGCGAGCTGTGTCACCGAACCCCGGAA 720
Db 983 AACTGTATATCTAAATATCCCAACAAAGGCGAGCTGTGTCACCGAACCCCGGAA 1042
QY 721 GAGTATCCAGAAAGTGCAGAAAGTGTATGAGAACTGAAAGGACCATATGCTGATCCCGGTG 780
Db 1043 GAGTATCCAGAAAGTGCAGAAAGTGTATGAGAACTGAAAGGACCATATGCTGATCCCGGTG 1102
QY 781 AGCATGTCTGAATCTGAAAGTGTGATGGTGTCTGCTACCTGCTGCTAGTTTAAATTAAC 840
Db 1103 AGCATGTCTGAATCTGAAAGTGTGATGGTGTCTGCTACCTGCTGCTAGTTTAAATTAAC 1162
QY 841 AAGAAGGTAGACTCTCTGA 858
Db 1163 AAGAAGGTAGACTCTCTGA 1180

RESULT 6	AX548046	AX548046	AX548046	4208 bp	DNA	linear	PAT 26-NOV-2002
LOCUS	AX548046	Sequence 22 from Patent WO02066654.					
DEFINITION	AX548046	AX548046					
ACCESSION	AX548046.1	GI:25813142					
VERSION							
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
FEATURES							
source							
ORIGIN							
Query Match							
Best Local Similarity							
Matches 854; Conservative							
0; Mismatches							
4; Indels							
0; Gaps							
1	ATGGCGCGGCTCGGGCACCCTTCGCTTCGGCGGGCCACCCACCGCTGTGTGGGGCG 60						
368	ATGGCGCGGCTCGGGCACCCTTCGCTTCGGCGGGCCACCCACCGCTGTGTGGGGCG 427						
61	CTACCCGAGTGTCTGTCCAGACGCGCTGAGAAGCGCCAAAGCGCGAGAGGTGCACGTC 120						
428	CTACCCGAGTGTCTGTCCAGACGCGCTGAGAAGCGCCAAAGCGCGAGAGGTGCACGTC 487						
121	GCCCGCGCGGAAACGGCAGCACACGCTCTACGTGGGCGTGTGGGCGAGCAAGCTGGGGCTG 180						
488	GCCCGCGCGGAAACGGCAGCACACGCTCTACGTGGGCGTGTGGGCGAGCAAGCTGGGGCTG 547						
181	CAGGTGGTGGAGCTGCGCGCGCGAGAGGCTTCCGAGCTGCTCTTCGTGAGAGACGCTG 240						
548	CAGGTGGTGGAGCTGCGCGCGCGAGAGGCTTCCGAGCTGCTCTTCGTGAGAGACGCTG 607						
241	GCCGTGGTGTGCGAGGAGACGGCCCTCATCACCGACCCCGGGCGCCGAGCGGAGGAG 300						
608	GCCGTGGTGTGCGAGGAGACGGCCCTCATCACCGACCCCGGGCGCCGAGCGGAGGAG 667						
301	GAGTTGACATGATGAAGAAGCATTAGAAAACCTCAGCTCAATATAGTAGAGATGAAA 360						
668	GAGTTGACATGATGAAGAAGCATTAGAAAACCTCAGCTCAATATAGTAGAGATGAAA 727						
361	GATGAATGCAACTTTAGATGGCGGAGATCTTTTATTACAGGCAGAGAAATTTTGTG 420						
728	GATGAATGCAACTTTAGATGGCGGAGATCTTTTATTACAGGCAGAGAAATTTTGTG 787						
421	GGCCTTCCAAAAGGACAAAATCAACGAGGTGCTGAAATCTTGGCTGATCTTTTAAGGAC 480						
788	GGCCTTCCAAAAGGACAAAATCAACGAGGTGCTGAAATCTTGGCTGATCTTTTAAGGAC 847						
481	TATGAGTCTCCACAGTCCAGTGGCAGATGGGTGCAATTTGAAGAGTCTTCTGCAGCATG 540						
848	TATGAGTCTCCACAGTCCAGTGGCAGATGGGTGCAATTTGAAGAGTCTTCTGCAGCATG 907						
541	GCTGGGCGCTTAACCTGATCGCAATTTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATC 600						
908	GCTGGGCGCTTAACCTGATCGCAATTTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATC 967						

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Boedet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ren Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasia van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAX Plate: 76 Row: j Column: 15.

Location/Qualifiers

FEATURES

source

1. .3737
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5295723"
/tissue_type="testis"
/clone_lib="NIH MGC_97"
/lab_host="DH10B"
/notes="Vector: pBluescript"
<1. .3737
/gene="DDAH1"
/note="synonym: DDAH"
/db_xref="LocusID:23576"
/db_xref="MIM:604743"
<1. .741
/gene="DDAH1"
/codon_start=1
/product="DDAH1 protein"
/protein_id="AAH43235.2"
/db_xref="GI:34783629"
/db_xref="LocusID:23576"
/db_xref="MIM:604743"
/translation="VARARQQLYVGLGSLQVVELPADSLPCVFVEDVAVV
CESTALITPGAPSRKREVMKREALEKLQNLNIVEMKBNALDGDVLFTRREFVVG
LSKRTNQRGAELIADTFDYAVSTVEVDGLHKGFCNAGNPLNIAIGSSSESAQALK
IMQMGSDHYRLKLTVPDDFAANCITLYNIKNHVLHRTPEYPSAKYKELKDHML
IPVMSLEKVDGLLTCCSVLNKKVDS"
1. .720
/gene="DDAH1"
/note="Aminotransf; Region: Amidinotransferase. This
family contains glycine (EC:2.1.4.1) and inosamine
(EC:2.1.4.2) amidinotransferases, enzymes involved in
creatine and streptomycin biosynthesis respectively. This
family also includes arginine deiminases, EC:3.5.3.6.
These enzymes catalyse the reaction: arginine + H2O <=>
citrulline + NH3. Also found in this family is the
Streptococcus anti tumor glycoprotein"
/db_xref="CDD:pfam02274"

gene

CDS

181 AAGGAGGTTGACATGATGAAAGAAGCATTAGAAAAAATCTCAGCTCAATATAGTAGATG 240
358 AAAGATCAAAATGCAACTTTTAGATGGCGGAGATGTTTTTATTCACAGGACAGAGATTTTTT 417
241 AAAGATCAAAATGCAACTTTTAGATGGCGGAGATGTTTTTATTCACAGGACAGAGATTTTTT 300
418 GTGGGCTTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTTGGTGTGATCTTTAAG 477
301 GTGGGCTTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTTGGTGTGATCTTTAAG 360
478 GACTATCGATCTCCACAGTCCGAGTCCGAGTGGTTCATTTGAAGAGTTTCTGCAGC 537
361 GACTATCGATCTCCACAGTCCGAGTGGTTCATTTGAAGAGTTTCTGCAGC 420
538 ATGGCTGGGCTTAACTCTGATGCGCAATTTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAG 597
421 ATGGCTGGGCTTAACTCTGATGCGCAATTTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAG 480
598 ATCATGCAACAGATGAGTGACCCCGCTACGACAACTCACTGTGCTGATGACATAGCA 657
481 ATCATGCAACAGATGAGTGACCCCGCTACGACAACTCACTGTGCTGATGACATAGCA 540
658 GCAAACTGTATATCTAAT 717
541 GCAAACTGTAT 600
718 GAAGATGATCCAGAAAGTGCAAGGTATATGAGAACTGAAGGACCATATGCTGATGATCCCC 777
601 GAAGATGATCCAGAAAGTGCAAGGTATATGAGAACTGAAGGACCATATGCTGATGATCCCC 660
778 GTGAGCATGCTGAACTGGAAAGGTGGATGGGTGCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCT 837
661 GTGAGCATGCTGAACTGGAAAGGTGGATGGGTGCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
838 AACAAAGAGGTAGACTCCTGA 858
721 AACAAAGAGGTAGACTCCTGA 741

RESULT 8
BC034505
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

BC034505 1120 bp mRNA linear ROD 25-NOV-2003
Mus musculus dimethylarginine dimethylaminohydrolase 1, mRNA (cdna
clone MGC:30267 IMAGE:5134831), complete cds.

BC034505 1120 bp mRNA linear ROD 25-NOV-2003
Mus musculus dimethylarginine dimethylaminohydrolase 1, mRNA (cdna
clone MGC:30267 IMAGE:5134831), complete cds.

BC034505 1120 bp mRNA linear ROD 25-NOV-2003
Mus musculus dimethylarginine dimethylaminohydrolase 1, mRNA (cdna
clone MGC:30267 IMAGE:5134831), complete cds.

BC034505 1120 bp mRNA linear ROD 25-NOV-2003
Mus musculus dimethylarginine dimethylaminohydrolase 1, mRNA (cdna
clone MGC:30267 IMAGE:5134831), complete cds.

BC034505 1120 bp mRNA linear ROD 25-NOV-2003
Mus musculus dimethylarginine dimethylaminohydrolase 1, mRNA (cdna
clone MGC:30267 IMAGE:5134831), complete cds.

BC034505 1120 bp mRNA linear ROD 25-NOV-2003
Mus musculus dimethylarginine dimethylaminohydrolase 1, mRNA (cdna
clone MGC:30267 IMAGE:5134831), complete cds.

TITLE

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

ORIGIN

Query Match 86.2%; Score 739.4; DB 9; Length 3737;
Best Local Similarity 99.9%; Pred. No. 2.8e-112;
Matches 740; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
118 GTGCGCGCGCGGAAAGCGAGACCGAGCTTACGTCGGCGGTGCTGGGAGCAGCTGGGG 177
1 GTGCGCGCGCGGAAAGCGAGACCGAGCTTACGTCGGCGGTGCTGGGAGCAGCTGGGG 60
178 CTGCGAGTGTGTGAGTGTGCGGCGCGGAGAGCTTCCGAGCTGCTTCGTCGAGGAGAC 237
61 CTGCGAGTGTGTGAGTGTGCGGCGCGGAGAGCTTCCGAGCTGCTTCGTCGAGGAGAC 120
238 GTGCGCGTGTGTGCGAGGAGACCGGCTTATCACCAGCCCGCGCGCGGAGCGGAGG 297
121 GTGCGCGTGTGTGCGAGGAGACCGGCTTATCACCAGCCCGCGCGGAGCGGAGG 180
298 AAGGAGGTTGACATGATGAAAGAAGCATTAGAAAAAATCTCAGCTCAATATAGTAGATG 357

MEDLINE
22388257
12477932
PUBMED
2 (bases 1 to 1120)
REFERENCE
Strausberg, R.
TITLE
Direct Submission
JOURNAL

Submitted (24-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: gcaps-x@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amc@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegue, H.,

Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,

A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 41 Row: 0 Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.

Location/Qualifiers

1. 1120

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:30267 IMAGE:5134831"

/tissue_type="Liver, normal. 5 month old male mouse."

/clone_lib="NCI CGAP L19"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. 1120

/gene="Ddahl1"

/note="synonym: 2510015N05Rik"

/db_xref="locusid:69219"

/db_xref="MGI:1916469"

122. 979

/codon_start=1

/product="dimethylarginine dimethylaminohydrolase 1"

/protein_id="AAH34505.1"

/db_xref="GI:21951624"

/db_xref="locusid:69219"

/translation="MAGLHPSAFGRATHAVVRAPESLCRHRLRSQEEVDFARAE
RQHELTVGLSGKLGVLQVLPADSLPDCVFVEDVAVVCEETALIRFGAPSRKEV
DMKEALEKLQNLIVEMDENLIDGDLVLTGREFVGLSKRTNORGAELIADTFK
YAVSTVPVADSLHLKSCNAGNLLIAGSSAQKALIKIMQMSDHRVTKLTPDDM
AANCYLNIIPSKGHVLLHRTPEEPESAKYVEXKLHLLIPVENSEMEKVDGLTCCS
VFINKKIDS"

155. 958

/note="Amidinotransf; Region: Amidinotransferase. This
family contains glycine (EC:2.1.1.4.1) and inosamine
(EC:2.1.1.4.2) amidinotransferases, enzymes involved in
creatine and streptomycin biosynthesis respectively. This
family also includes arginine deiminases, EC:3.5.3.6.
These enzymes catalyse the reaction: arginine + H2O <=>
citrulline + NH3. Also found in this family is the
Streptococcus anti tumour glycoprotein"

/db_xref="CDD:pfam02274"

ORIGIN

Query Match 83.6%; Score 717.2; DB 10; Length 1120;

Best Local Similarity 89.7%; Pred No. 1.3e-108;

Matches 770; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY	1	ATGGCCGGCCTCGGCCACCCCTCGCCTTGGCGGGGCCACCCACGCGGTGCTGGGGCG	60
DB	122	ATGGCCGGCCTCGGCCACCCCTCGCCTTGGCGGGGCCACCCACGCGGTGCTGGGGCG	181
QY	61	CTATCCCGAGTCTGTCACGACGCGCTGAGAGCGCCCAAGGGCGAGAGGTGAGATC	120
DB	182	CGGCCGAGTCCCTGTGCGGCCACGCGCTGAGGCGCTCGCAGGGCGAGGAGTGAATTC	241
QY	121	GCCCGCGGGAACGGCAGCACACAGCTCTAGTGGCGCTGCTGGCAGCAGCAGCTGGGGTG	180
DB	242	GCTCGCGCGGAGCGCCAGCAGAGCTCTAGTGGCGCTGCTGGCAGCAGCAGCTGGGGTG	301
QY	181	CAGTGTGTGAGTCTGCGCGCGCAGCAGAGCCTTCGCGACTCGCTTCGTGAGAGAGT	240
DB	302	CAGTGTGTGAGTCTGCGCGCGCAGCAGAGCCTTCGCGACTCGCTTCGTGAGAGAGT	361
QY	241	GCCGTGTGTGCGAGGAGACGGCCCTCATCACCGACCCCGGGCGCGCAGCGAGGAG	300
DB	362	GCCGTGTGTGCGAGGAGACGGCCCTCATCACCGCGCGCGGGCGCGCAGCGAGGAG	421
QY	301	GAGGTGTGACATGATGAAGAAGCAATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAA	360
DB	422	GAGGTGTGACATGATGAAGAAGCAATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAA	481
QY	361	GATGAAATGCACTTTAGATGGCGGAGATGTTTTATTCACAGCGAGAGATTTTTTGTG	420
DB	482	GATGAAATGCACTTTGATGGTGGGACGCTCCTATTTCACAGCGAGAGATTTTTTGTG	541
QY	421	GGCTTTCCAAAAGGACAAATCAACGAGGTGCTCAAAATCTTTGGCTGATACATTTAAAGAC	480
DB	542	GGCTTTCCAAAAGGACAAATCAACGAGGTGCTCAAAATCTTTGGCTGATACATTTAAAGAC	601
QY	481	TATGCAGTCTCCACAGTCCAGTGGCAGATGGTGTGCAATTTGAAGAGTTCTGCAGCATG	540
DB	602	TACGCAGTCTCTACAGTCCCTGTGCGCGATTTCTTTGCAATTTAAAGAGTTCTGCAGCATG	661
QY	541	GCTGGCTTAACCTGATCGCAATTTGGTCTAGTGAATCTGCACAGAGGCGCTTTAAGATC	600
DB	662	GCGGACCCCACTGATTCGAATAGGTCACGCAATCTGCACAGAGGCGCTTTAAGATC	721
QY	601	ATGCAACAGATGAGTGACCCGCTACGACAACTCACTGTGCTGTGATGACATAGCAGCA	660
DB	722	ATGCAACAGATGAGTGACCCGCTATGCAAGCTCACTGTACCCGACGACATGCGCCGC	781
QY	661	AACTGTATATCTTAATATATCCCAAGAGGACGCTTTGCTGACCGACGACCCCGGAA	720
DB	782	AACTGTATATCTTAATATATCCCAAGAGGACGCTTTGCTGACCGACGACCCCGGAA	841
QY	721	GAGTATCCAGAAAGTGCAAGGTTTATGAGAAACTGAGAGGACCATATGCTGATCCCGTG	780
DB	842	GAGTATCCAGAAAGTGCAAGGTTTATGAGAAACTGAGAGGACCATATGCTGATCCCGTG	901
QY	781	AGCATGTCTGAACCTGGAAGAGTGGATGGGCTGCTACCTGCTGCTGATTTAATTAAC	840
DB	902	AGCAACTCGAGATGGAAGAGTGGATGGGCTGCTACCTGCTGCTGATTTAATTAAC	961
QY	841	AGAAAGTGTAGTCTCTGA 858	
DB	962	AGAAAGTGTAGTCTCTGA 979	

RESULT 9

LOCUS

AX683109

DEFINITION

Sequence 83 from Patent EPI279744.

ACCESSION

AX683109

VERSION

AX683109.1

GI:29370146

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

linear

3008 bp

DNA

Sequence 83 from Patent EPI279744.

AX683109

AX683109.1

GI:29370146

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

linear

PAT 29-MAR-2003

Rattus.
1
Brooksbank, R.A., Dixon, A.K., Lee, K. and Pinnock, R.D.
Identification and use of molecules implicated in pain
Patent: EP 1279744-A 83 29-JAN-2003;
WARNER-LAMBERT COMPANY (US)
JOURNAL
FEATURES
source
Location/Qualifiers
1..3008
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
/note="N-G, N-G dimethylarginine dimethylaminohydrolase"

Query Match 83.4%; Score 715.6; DB 6; Length 3008;
Best Local Similarity 89.6%; Pred. No. 2.4e-108;
Matches 769; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 1 ATGGCGGGCTCGGCGACACCCCTCGCGCTTCGGCGGGGCCACCGCGCTGGTGGCGG 60
DB 432 ATGGCGGGCTCAGCACCCCTCGCTTCGGCGGGGCCACCGCGCTGGTGGCGGCT 491

QY 61 CTACCCGAGTCTGCGACGACGCTGAGAGCGGCGGCGGCGGCGGCGGCGGCGG 120
DB 492 CGCGCGGAGTCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 551

QY 121 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
DB 552 GTCGCGGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 611

QY 181 CAGGTGGTGGAGTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
DB 612 CAGGTGGTGGAGTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 671

QY 241 GCGGTGGTGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 672 GCGGTGGTGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 731

QY 301 GAGGTGGATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 732 GAGGTGGATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 791

QY 361 GATGAAATGCAATTTAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
DB 792 GATGAAATGCAATTTAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 851

QY 421 GCGCTTTCCAAAGGAGCAATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
DB 852 GCGCTTTCCAAAGGAGCAATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 911

QY 481 TATGAGTCTCCACAGTGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
DB 912 TATGAGTCTCCACAGTGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 971

QY 541 GCTGGGCTTAAGTCTGATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
DB 972 GCTGGGCTTAAGTCTGATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1031

QY 601 ATGCAACAGATGAGTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
DB 1032 ATGCAACAGATGAGTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1091

QY 661 AACTGTATATATCTAATATCTCCCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
DB 1092 AACTGTATATATCTAATATCTCCCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1151

QY 721 GAGTATCCAGAAAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
DB 1152 GAGTATCCAGAAAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1211

QY 781 AGCATGTCTGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
DB 1212 AGCATGTCTGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1271

QY 841 AAGAAGGTAGCTCTCTGA 858
DB 1272 AAGAAGGTAGCTCTCTGA 1289

RESULT 10
D86041
LOCUS
DEFINITION
Rattus norvegicus mRNA for N-G,N-G-dimethylarginine dimethylaminohydrolase, complete cds.
ACCESSION
D86041
VERSION
D86041.1 GI:1906799
KEYWORDS
N-G,N-G-dimethylarginine dimethylaminohydrolase.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE
1 (sites)
Kimoto, M., Sasakawa, T., Tsuji, H., Miyatake, S., Oka, T., Nio, N. and Ogawa, T.
Cloning and sequencing of cDNA encoding NG, NG-dimethylarginine dimethylaminohydrolase from rat kidney
Biochim. Biophys. Acta 1337 (1), 6-10 (1997)
MEDLINE
97157050
PUBMED
9003431
REFERENCE
2 (bases 1 to 3008)
Kimoto, M.
AUTHORS
Cloning and sequencing of cDNA encoding G-N, G-N-dimethylarginine dimethylaminohydrolase from rat kidney
Unpublished
REFERENCE
3 (bases 1 to 3008)
Kimoto, M.
AUTHORS
Direct Submission
TITLE
Submitted (14-JUN-1996) Masumi Kimoto, The University of Tokushima, School of Medicine, Nutrition; Kuramoto 3, Tokushima, Tokushima 770, Japan (E-mail: kimoto@nutr.med.tokushima-u.ac.jp, Tel: 0886-33-7088, Fax: 0886-33-7089)
JOURNAL
FEATURES
source
Location/Qualifiers
1..3008
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Wistar"
/db_xref="taxon:10116"
/sex="female"
/tissue type="kidney"
432..1289
/note="DDAH"
/codon_start=1
/product="N-G,N-G-dimethylarginine dimethylaminohydrolase"
/protein_id="BAA18993.1"
/db_xref="GI:1906800"
/translation="MAGLSHPVFGRAHVRAPPESLCRHARRSQGEVDFARAE ROHLYVGLGSLGVQLPADESLPCVFVEDVAVVCEETALITRPGAPRRKEV DMKALEKQLNIEMKDNATLDGGDLFTGREFFVGLSKETNORGAELIADTFKD YAVSTVPADSLHLKFCFSGMAGNLIAISSSAQKALIKMQMSBHRVYDKLTVPDDM AAVCIYINIPSKGHVILLHTPPEYPSAKYKLEKDLHLLIPVSNSEMEKVDGLLTCCS VFINKKTDs"

Query Match 83.4%; Score 715.6; DB 10; Length 3008;
Best Local Similarity 89.6%; Pred. No. 2.4e-108;
Matches 769; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 1 ATGGCGGGCTCGGCGACACCCCTCGCGCTTCGGCGGGGCCACCGCGCTGGTGGCGG 60
DB 432 ATGGCGGGCTCAGCACCCCTCGCTTCGGCGGGGCCACCGCGCTGGTGGCGGCT 491

QY 61 CTACCCGAGTCTGCGACGACGCTGAGAGCGGCGGCGGCGGCGGCGGCGGCGG 120
DB 492 CGCGCGGAGTCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 551

121 GCCCGCGGAACCGCAGCAGCTCTAGTGGGCGTCTGGGCGAGCAGCTGGGCGTG 180
Db GCTCGCGCTGAGCCAGCAGCAGCTCTAGTGGGCGTCTGGGCGAGCAGCTGGGCGTG 611
Qy CAGTGTGTGAGCTGCGCGCGCAGCAGCAGCTCTGGGCGAGCAGCTGGGCGTG 240
Db CAGTGTGTGAGCTGCGCGCGCAGCAGCAGCTCTGGGCGAGCAGCTGGGCGTG 671
Qy GCCGTGTGTGCGAGGAGCGCGCTCATCACCGACCGCGCGCGCGCGCGAGGAG 300
Db GCCGTGTGTGCGAGGAGCGCGCTCATCACCGACCGCGCGCGCGCGCGAGGAG 731
Qy GAGTGTGACATGATGAAGAGCAATTAAGAACTTCAGCTCAATATAGTAGAGTGA 360
Db GAGTGTGACATGATGAAGAGCAATTAAGAACTTCAGCTCAATATAGTAGAGTGA 791
Qy GATGAATGCACTTTAGTGGCGAGATGTTTATTCACAGCAGAGATTTTGTG 420
Db GATGAATGCACTTTAGTGGCGAGATGTTTATTCACAGCAGAGATTTTGTG 851
Qy GGCCTTTCCAAAGAGCAATTAAGAACTTCAGCTCAATATAGTAGAGTGA 480
Db GGCCTTTCCAAAGAGCAATTAAGAACTTCAGCTCAATATAGTAGAGTGA 911
Qy TATGAGTCTTCCAGTGCAGTGGCAGATGGGTGATTTGAAGATTTCTGCGAGATG 540
Db TATGAGTCTTCCAGTGCAGTGGCAGATGGGTGATTTGAAGATTTCTGCGAGATG 971
Qy GCTGGCCTACCTGATCGCAATTTGGGTAGTGAATTCACAGCAGAGCGCTTATGATC 600
Db GCTGGCCTACCTGATCGCAATTTGGGTAGTGAATTCACAGCAGAGCGCTTATGATC 1031
Qy ATGCAACAGATGAGTACCGCTATGACAACTCTGCTGCTGATGATGATGATGATG 660
Db ATGCAACAGATGAGTACCGCTATGACAACTCTGCTGCTGATGATGATGATGATG 1091
Qy AACTGTATATCTAAATATCCCAACAGAGGACAGCTTCTGCTGACCGACCGCGAA 720
Db AACTGTATATCTAAATATCCCAACAGAGGACAGCTTCTGCTGACCGACCGCGAA 1151
Qy GAGTATCCAGAAAGTGAAGGTTTATGAGAACTGAAGACCATATGCTGATCCCGTG 780
Db GAGTATCCAGAAAGTGAAGGTTTATGAGAACTGAAGACCATATGCTGATCCCGTG 1211
Qy AGCATGCTGAACGTGAAAAGTGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db AGCATGCTGAACGTGAAAAGTGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1271
Qy 841 AAGAGGTAGACTCTGA 859
Db 1272 AAGAGACAGACTCTGA 1289

RESULT 11
AB095027
LOCUS AB095027
DEFINITION Gallus gallus mRNA for dimethylarginine dimethylaminohydrolase I,
complete cds.
ACCESSION AB095027
VERSION AB095027.2
KEYWORDS Gi:32306458
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1
AUTHORS Mishima, T., Hamada, T., Ui-Tei, K., Takahashi, F., Miyata, Y.,
Imaki, J., Suzuki, H. and Yamashita, K.
TITLE The expression of DDH I in the chick and rat embryo
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1150)
AUTHORS Mishima, T., Hamada, T., Ui-Tei, K., Takahashi, F., Miyata, Y.,

Imaki, J., Suzuki, H. and Yamashita, K.
Direct Submission
Submitted (30-OCT-2002) Takuya Mishima, Nippon Medical School,
Department of Anatomy, Sendagi 1-1-5, Bunkyo, Tokyo 113-8602, Japan
(E-mail: mishima@nms.ac.jp, Tel: 81-3-3822-2131 (ex. 5315))
On Jun 26, 2003 this sequence version replaced gi:30268647.
COMMENT Location/Qualifiers
FEATURES
1..1150
source
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/tissue_type="embryos yolk sac"
/clone_lib="Chick Yolk Sac Lambda cDNA Library (Stratagene
#937406)"
/dev_stage="Five-days-old chick"
40..903
codon_start=1
product="dimethylarginine dimethylaminohydrolase I"
protein_id="BAC7602.1"
/db_xref="GI:30268648"
translation="MAGLGGPFAAFGRCTHAAVRALPESLRQALRSTAGPEVDFA
REHKLVTGLRGLQVLELPADSELPDCVFVEDAAVCEETALLRPGAPRRKE
VEAKHVLNLANVMDENALDGGDLFTGREFEVLGSKTNGRGAELLADTF
KDYAVTFVHDSLRKFCSMAGFNLAIGSSENAQAALKTQMOMSHRDKLTPD
DAACIYINIPSKGHVLLHRAPEYPESAKVFEKLKDHMLPIPIANTELEKVDGALTC
CSVLINKTSEL"
ORIGIN
Query Match 67.6%; Score 579.6; DB 5; Length 1150;
Best Local Similarity 80.9%; Pred. No. 6.4e-86;
Matches 688; Conservative 0; Mismatches 159; Indels 3; Gaps 1;
QY 12 CGGCCACCCCTCCGCTTGGCGGCGCCACCCACCCCTGGTGGCGGCTACCCGAGTC 71
Db 54 CGCGCGCCCGCGCGCTTGGCGGCGCTGTACGACCGCTGGTGGCGGCTGCCGAGTC 113
QY 72 GCTCTGCCAGCAGCGCTGAGAGCGCCAGCGGCGAGGAGTGGACGTCGCCGCGCGGA 131
Db 114 GCTGTGCCCGCAGGCGTTCGCGACGCGCGCGCGCGCGCGCGCTTCGCTCGCGCGGA 173
QY 132 AGCGCAGCACCGCTCTAGTGGCGGCTGCTGGGCGCAAGCTGGGCGCTGCAGTGGTGA 191
Db 174 GCGGGAGCATCAGTGTACGTGGCGGTCTACGCGGCAAGCTGGGCGCTGCAGTGGTGA 233
QY 192 GCTGCGCGCGCAGAGAGCTTCGCGACGCTGCTTCGTTGGAGGAGCTGGCGGCTGTG 251
Db 234 GCTGCGCGCTGACGAGAGCTCCCGGCTGCTTCGTTGGAGGAGCTGGCGGCTGTG 293
QY 252 CGAGGAGACGCGCTCATCACCGCGCGCGCGCGCGCGCGCGCGCGCGAGAGGAGTGA 311
Db 294 CGAGGAGACGCGCTGCTCACCGCGCGCGCGCGCGCGCGCGCGCGCGAGAGGAGTGA 353
QY 312 GATGAAGAGCATTTAGAAAACTTCAGCTCAATATAGTAGAGTGAAGATGAAGATGC 371
Db 354 CATGAAGAGAGTACTAGAAAGCTTTAACTCAACATGGTAGAGATGGTAGAGATGC 413
QY 372 AACTTTAGATGCGGAGATGTTTATTCACAGGCGAGAGAA---TTTTTTGGGCTTTC 428
Db 414 AACTTTGATGTTGAGAGAGCTTTATTTACAGGCGAGAGATTTTGTAGGCGCTTTC 473
QY 429 CAAAAGGCAAAATCAACGAGGTGCTGAATCTTGCTGATCTTTTAAAGACTATGAGT 488
Db 474 CAAGCGCAAAATCAACGCGTGGGCGAGAAATCTGGCTGACACATTTAAGGATTTATGCT 533
QY 489 CTCACAGTGGCAGTGGCAGATGGCTTGCATTTGAAGAGTTTCTCAGCATGCTGGGCG 548
Db 534 CTCACAGTGGCAGTGGCAGATGGCTTGCATTTGAAGAGTTTCTCAGCATGCTGGGCG 593
QY 549 TAACTGTATGCGAATTTGGTCTAGTGAATCTTCACAGAGGCGCTTAAAGATCATGCAACA 608
Db 594 AAACTGTATGCTATCGGCTCAAGTGAAGTGCACAGAAAGGCTTCAAGACCATCAACA 653
QY 609 GATGAGTACACCGCTAGCAGAAACTCTGCTGCTGATGATGATGATGATGATGAT 668

Db	654	GATGACGACCCACCGCTACGACAACTGACGGTGCCGACGATGCCGCCCAACTGCAT	713
QY	669	ATATCTAAATATCCCCAACAAAGGCGACGCTTCTGCTGCACCGAACCCCGGAAGATATCC	728
Db	714	CTACTTAAACATTTCCACGACAAAGGCGACGCTCTGCTGCACCGAGCCCTCAGGAGTACCC	773
QY	729	AGAAAGTCGAAGGTTTATGAGAACTGAGAGGACCATGCTCATCCCCGTGAGCATGTC	788
Db	774	AGAGAGCGCAAGGTTTTTGAAGAACTGAAGGACCATGCTCATCCCCATAGCCAAAC	833
QY	789	TGAAGTGGAAAAAGTGGATGGGTGCTCACCTGCTGCTCAGTTTAAATTAACAAGAGGT	848
Db	834	AGAACTGGAGAAAGTAGATGGGCACTCACCTGCTGCTGCTTATTAAACAAACTTC	893
QY	849	AGACTCCTGA	858
Db	894	AGAAATATGA	903
RESULT 12	HSM808292	3908 bp mRNA linear	PRI 30-AUG-2003
LOCUS	Homo sapiens mRNA; cDNA DKFZp686N2176 (from clone DKFZp686N2176).		
DEFINITION	EX648145		
ACCESSION	EX648145.1	GI:34367304	
VERSION			
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 3908) Ansoez,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B., Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.		
CONSRMT	The German Human cDNA Consortium		
TITLE	Direct Submission		
JOURNAL	Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764		
COMMENT	Neuhberg, GERMANY Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de; sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp686N2176) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/ .		
FEATURES	Location/Qualifiers		
source	1..3908 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="DKFZp686N2176" /tissue_type="human endometrium carcinoma cell line" /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host DH10B; sites SfiIA + SfiIB" /dev_stage="adult"		
ORIGIN			
Query Match	84.7%;	Score 555.4;	DB 9; Length 3908;
Best Local Similarity	99.8%;	Pred. No. 6.1e-82;	
Matches 556; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
QY	302	AGCTTGACATGATGAAGAAGCATTAGAAAACTTCAGCTCAATATAGTAGATGAAG	361
Db	382	AGCTTGACATGATGAAGAAGCATTAGAAAACTTCAGCTCAATATAGTAGATGAAG	441
QY	362	ATGAAAATGCAACTTTAGATGGCGAGATGTTTTTATTCACAGGAGAAATTTTTGTGG	421
Db	442	ATGAAAATGCAACTTTAGATGGCGAGATGTTTTTATTCACAGGAGAAATTTTTGTGG	501

[illegible][illegible]

882 GTTGATGCTGTGACCTGCTGCTCCATCCTTATTAAACAAAGAAAGTCTCTGA 938

RESULT 15

BC060266	BC060266	3420 bp	mRNA	linear	ROD 25-NOV-2003
LOCUS					
DEFINITION	Mus musculus dimethylarginine dimethylaminohydrolase 1, mRNA (cdna clone IMAGE:5698949), complete cds.				

REFERENCE
AUTHORS
1 (bases 1 to 3420)
Srausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodríguez, S.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,

TITLE Generation and initial analysis of more than 15,000 full-length

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932 PUBMED

AUTHORS
Strausberg, R.
Direct Submission

JOURNAL
Submitted (20-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC) Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA

REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT	Contact: MGC help desk

Email: cgapbs-remail.hhh.gov
 tissue procurement: Dr. Jim Lin. University of Iowa

CDNA Library Preparation: M. Benito Soares, University of Iowa
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada

INFO@CYS.CC.CA
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,

Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo

Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline

Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,

THE UNIVERSITY OF CHICAGO

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

This clone was selected for full length sequencing because it

analysis

the longest cds in the locus.

Search completed: June 7, 2004, 20:25:47
Job time : 3417 secs

```
source      1. .3420
             /organism="Mus musculus"
             /mol_type="mRNA"
             /strain="C57BL/6"
             /db_xref="taxon:10090"
             /clone="IMAGE:5698949"
             /tissue_type="Brain, whole, 18.5 dpc mouse embryo"
             /clone_lib="NIH_EMAP_EQO"
             /lab_host="DH10B"
             /note="Vector: pYX-ASC"
             1. .3420
             /gene="Ddahl1"
             /note="synonym: 2510015N06Rik"
             /db_xref="LocusID:62219"
             /db_xref="MGI:1916469"
             360. .680
             /codon_start=1
             /product="Ddahl1 protein"
             /protein_id="AAH60266.1"
             /db_xref="GI:37805214"
             /db_xref="LocusID:62219"
             /translation="MAGFNLRIGSSSAQKALKINQMSDHRDKLTPDDMAANCI
             YLNPSPKGVHLLHRTPEEPESAKVYEKLDHLLIPVNSEMEKVDGLLTCCSVFINK
             KIDS"
             360. .659
             /note="Aminotransf; Region: Amidinotransferase. This
             family contains glycine (EC:2.1.4.1) and inosamine
             (EC:2.1.4.2) amidinotransferases, enzymes involved in
             creatine and streptomycin biosynthesis respectively. This
             family also includes arginine deiminases, EC:3.5.3.6.
             These enzymes catalyze the reaction: arginine + H2O <=>
             citrulline + NH3. Also found in this family is the
             Streptococcus anti tumour glycoprotein"
             /db_xref="CDD:pfam02274"

misc_feature
             360. .659
             /note="Aminotransf; Region: Amidinotransferase. This
             family contains glycine (EC:2.1.4.1) and inosamine
             (EC:2.1.4.2) amidinotransferases, enzymes involved in
             creatine and streptomycin biosynthesis respectively. This
             family also includes arginine deiminases, EC:3.5.3.6.
             These enzymes catalyze the reaction: arginine + H2O <=>
             citrulline + NH3. Also found in this family is the
             Streptococcus anti tumour glycoprotein"
             /db_xref="CDD:pfam02274"

ORIGIN
Query Match      44.5%; Score 381.8; DB 10; Length 3420;
Best Local Similarity 88.8%; Pred. No. 2.5e-53;
Matches 413; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 394 TTATTTCACAGGCAGAGAAATTTTGTGGGCTTTCCAAAAGGACAAATCAACGAGGTGCT 453
DB 216 TTGCTCAGCAGCAGAGAAATTTTGTGGGCTTTCCAAAAGGACAAATCAACGAGGTGCT 275
QY 454 GAAATCTTGGCTGATCTTTAAGCACTATGCACTCTCCACAGTGCAGTGGCAGATGGG 513
DB 276 GAAATCTTGGCTGATCTTTAAGCACTATGCACTCTCCACAGTGCAGTGGCAGATGGG 335
QY 514 TTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCTTAACCTGATCGCAATTGGGTCTAGT 573
DB 336 TTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCTTAACCTGATTCGAATAGGGTCCAGC 395
QY 574 GAAATCTGCACAGAGGCGCTTTAAGATCATGCAACAGATGAGTGACCAACCGCTACGACAA 633
DB 396 GAAATCTGCACAGAGGCGCTTCAGATCATGCAACAGATGAGTGACCAACCGCTATGACAAG 455
QY 634 CTCAGTGTGCTGATGATAGCAGCAAACTGTATATATCTAATATATCCCAACAAAGGG 693
DB 456 CTCAGTGTACCCGACGATGCGCGCACTGCAATATATCTAATATATCCCAACAAAGGG 515
QY 694 CACGTCTTGTGTCACCGAACCCCGAGAGATATCCAGAAAGTGCAAAAGTTTATGAGAAA 753
DB 516 CATGCTGTGTCACCGAACCCCGAGAGATATCCAGAAAGGCGCAAGGTCATGAGAAA 575
QY 754 CTGAAGGACCAATATGTGATCCCGTGAGCATGTCTGAACTGGAAAAGGTGGATGGGCTG 813
DB 576 CTCGAAGGACCAATCTACTGATCCCTGTGAGCAACTCGGAGATGGAAAAGGTGGACGGCTTG 635
QY 814 CTCACCTGCTGCTCAGTTTAAATTAACAAGAGGTAGACTCTCTGA 858
DB 636 CTCACCTGCTGCTCAGTTTAAATTAACAAGAGGTAGACTCTCTGA 680
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M protein - nucleic search, using frame_plus_p2n model

run on: June 7, 2004, 22:28:56 ; Search time 3208 Seconds
(without alignments)
3850.612 Million cell updates/sec

title: US-09-889-733B-2
perfect score: 1454
sequence: 1 MAGLHPSAFGRATHAVVRA.....EKVDGLITCSVLINKKVD 285

scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

searched: 3470272 seqs, 21671516995 residues

total number of hits satisfying chosen parameters: 6940544

minimum DB seq length: 0
maximum DB seq length: 2000000000

post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

command line parameters:
MODEL=frame+ p2n.model -DEV=xlp
Q=/cgn2_1/USPTO.spool.p/US09889733/runat_07062004_083528_7970/app_query.fasta_1.455
DB=GenEmbl -QMT=fastp -SUPFIX=p2n.rge -MINMATCH=0.1 -LOPCL=0 -LOOPXT=0
UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
USER=US09889733 @CGN 1 1 5600 @runat_07062004_083528_7970 -NCPU=3
NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_im:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_ats:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_on:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_ats:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rdd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	1454	100.0	858	6	BD251986 Screen me
2	1454	100.0	858	6	AX032835 Sequence
3	1439	99.0	1624	9	BC033680 Homo sapi
4	1439	99.0	1633	6	BD094073 Shear str
5	1439	99.0	1633	9	AB001915 Homo sapi
6	1439	99.0	4208	6	AX548046 Sequence
7	1389	95.5	1120	10	BC034505 Mus muscu
8	1379	94.8	3008	6	AX683109 Sequence
9	1379	94.8	3008	10	D86041 Rattus norv
10	1253	86.2	3737	9	BC043235 Homo sapi
11	1243	85.5	1150	5	AB095027 Gallus ga
12	1149.5	79.1	1344	5	BC056074 Xenopus l
13	941	64.7	3908	9	BSM808292 Homo sapi
14	910	62.6	2289	10	AX128909 Mus muscu
15	756	52.0	3420	10	BC060266 Mus muscu
16	696	47.9	1310	10	BC003328 Mus muscu
17	694	47.7	1263	9	BC001435 Homo sapi
18	694	47.7	1351	6	AK098671 Homo sapi
19	694	47.7	1351	9	AX779950 Sequence
20	694	47.7	1351	9	AF087894 Homo sapi
21	694	47.7	1376	6	AX011712 Sequence
22	694	47.7	1376	6	BD226323 Pancrati
23	694	47.7	1709	9	AK026191 Homo sapi
24	688	47.3	858	6	BD251987 Screen me
25	688	47.3	858	6	AX032837 Sequence
26	661.5	45.5	1270	10	AF004106 Mus muscu
27	645.5	44.4	1397	6	AX379492 Sequence
28	624	42.9	1228	6	AX780306 Sequence
29	546	37.6	2895	9	AK098405 Homo sapi
30	502	34.5	104071	2	AL162260 Homo sapi
31	502	34.5	158405	2	AC019261 Homo sapi
32	502	34.5	168953	9	AL360219 Human DNA
33	483	33.2	186434	2	AC126412 Mus muscu
34	482	33.1	1377	3	AY060959 Drosophil
35	476	32.7	228940	2	AC126890 Rattus no
36	474	32.6	135545	10	MHC213L3 AF109905 Mus muscu
37	474	32.6	221893	10	AC087117 Mus muscu
38	449.5	30.9	228453	2	AC094348 Rattus no
39	423	29.1	28772	9	HSA012008 AL04216 Human sapi
40	423	29.1	53125	9	AL844216 Human DNA
41	423	29.1	54146	9	EX248244 Human DNA
42	423	29.1	90732	9	AL670886 Human DNA
43	423	29.1	100000	9	AP000503 Homo sapi
44	423	29.1	136493	9	AL662899 Human DNA
45	423	29.1	163662	9	AL132713 Human DNA

ALIGNMENTS

RESULT 1

BD251986	BD251986	858 bp	DNA	linear	PAT 17-JUL-2000
LOCUS	Screen method.				
DEFINITION	BD251986				
ACCESSION	BD251986				
VERSION	BD251986.1	GI:33061756			
KEYWORDS	JP 2002535001-A/1.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 858)				
TITLE	Vallance,P.J.T., Leiper,J.M., Whitley,G.S.J. and Charles,I.G.				
JOURNAL	Screen method				
COMMENT	Patent: JP 2002535001-A 1 22-OCT-2002; UNIVERSITY COLLEGE LONDON				
	OS Homo sapiens (human)				
	PN JP 2002535001-A/1				
	PD 22-OCT-2002				
	PF 26-JAN-2000 JP 2000596131				
	PR 26-JAN-1999 GB 9901705.5 04-JUN-1999 GB 9913066.8 PI				
	PATRICK JOHN THOMPSON VALLANCE,JAMES MITCHELL LEIPER,GUY ST PI				
	JOHN WHITLEY,				
	PI IAN GEORGE CHARLES				
	PC C12N15/09,A01K67/027,A61K31/198,A61K38/00,A61K45/00,A61P3/06, PC				
	PC A61P9/02,				
	PC A61P9/04,A61P9/10,A61P9/12,A61P13/12,A61P25/06,A61P25/18, PC				
	A61P25/28				
	PC A61P29/00,A61P31/04,A61P35/00,C07K16/40,C12N1/15,C12N1/19, PC				
	C12N1/21,				
	PC C12N5/10,C12N9/78,C12Q1/02,C12Q1/34,G01N33/15,G01N33/50 PC				
	,C12N15/00,C12N5/00,				
	PC A61K37/02				
	CC Screen method				
	PH Key				
	FT CDS				
FEATURES	Location/Qualifiers				
source	1..858				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
ORIGIN	(1)..(858).				
Alignment Scores:					
Pred. No.:	6.53e-124	Length:	858		
Score:	1454.00	Matches:	285		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-09-889-733B-2 (1-285) x BD251986 (1-858)					
QY	1	MetaAGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla	20		
Db	1	ATGGCGCGCCTCGCCACCCCTCGCCTCGCGGGCCACCACGCCCTGGTGGGGCG	60		
QY	21	LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluValAspVal	40		
Db	61	CTACCCGAGTCGCTCGCCAGCAGCGCTGAGAAGGCCCAAGGCGGAGGAGCGTGC	120		
QY	41	AlaArgAlaGluArgGlnHisGlnLeuTyValGlyValLeuGlySerLysLeuGlyLeu	60		
Db	121	GCCGCGCGGAAACGGCAGCACCAGCTCTAGTGGGCGTCTGGGCGAAGCTGGGGCTG	180		
QY	61	GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal	80		
Db	181	CAGTGGTGGAGCTCGCGGCGGAGCAGAGAGCTTCGAGACTTCCGACTTGCCTTC	240		
QY	81	AlaValValCysGluGluThrAlaLeuIleThrArgProGlyValaProSerArgGlys	100		
Db	241	GCCCTGGTGTCCGAGGAGACGGCCCTCATCACCGACCCCGGGCGCCGAGCGGAGG	300		
QY	101	GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys	120		

VLINKKVS"

ORIGIN

Alignment Scores: 6 53e-124 Length: 858
Pred. No.: 1454.00 Matches: 285
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6

JS-09-889-733B-2 (1-285) x AX032835 (1-858)

2Y 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
Db 1 ATGGCGCGGCTCGGCCACCCCTTCGCGCTTCGCGCGGCCACCCACCGCTGTCGGGCGG 60
2Y 21 LeuProGluSerLeuGlyHisAlaLeuArgSerAlaValGlyGluValValArgVal 40
Db 61 CTACCCGAGTCGCTTCGCCAGACCGCGCTGAGAAGGCCAGCGGAGGAGTGACGTC 120
2Y 41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLeuGlyLeu 60
Db 121 GCCCGCGCGGAGCGGAGCAGCACCGCTCTACGTGGGCGTCTCGGCGAGCAAGCTGGGCGTG 180
2Y 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
Db 181 CAGGTGGTGGAGTCGCGCGCGGAGGAGCGCTTCGCGACTGCGTCTTCGTGGAGGACGTCG 240
2Y 81 AlaValValCysGluGluThrAlaLeuLeuThrArgProGluValAlaProSerArgArgLys 100
Db 241 GCCGTGGTGGGAGGAGCGGCGCTCATCCCCGCGCGCGCGCGCGCGGAGGAG 300
2Y 101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
Db 301 GAGGTTCGATGATGAAGAAGCATTAGAAAAAATTCAGCTCAATATAGTAGAGATGAAA 360
2Y 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal 140
Db 361 GATGAAAATGCACTTTAGATGCGGAGATGTTTATTCACAGCGAGAGATTTTGTG 420
2Y 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuAlaAspThrPheLysAsp 160
Db 421 GGCGTTTCCAAAAGGAGCAAAATCAAGAGGTGCTGAAATCTTCGCTGATACTTTAAAGGAC 480
2Y 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
Db 481 TATGCGAGTCTCCACAGTCGCGAGTGGGAGATGGTTCGATTTGAAGAGTTTCGCGAGCATG 540
2Y 181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200
Db 541 GCTGGGCGCTTAACCTGATCGCAATTCGGTCTAGTGAATCTGCACAGAGCCCTTAAGATC 600
2Y 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
Db 601 ATGCAACAGATGAGTGACCCAGCGCTACGCAAAACTCACTGTGCGCTGATGACATAGCA 660
2Y 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
Db 661 ACTGTATATATCTAATATATCCCAACAAAGGCGACGCTTCTGCTGACCGAACCCCGGAA 720
2Y 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal 260
Db 721 GAGTATCCAGAAAGTGCAAGGTTTATGAGAACTGAAGGACCATATGCTGATCCCGCTG 780
2Y 261 SerMetSerGluLeuGluLysValAspGlyLeuThrCysCysSerValLeuIleAsn 280
Db 781 AGCATGCTCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
2Y 281 LysLysValAspSer 285
Db 841 AAGAAGGTAGACTCC 855

RESULT 3
BC033680
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE

AUTHORS
TITLE
JOURNAL

REMARK
COMMENT

BC033680 1624 bp mRNA linear PRI 07-OCT-2003
Clone MGC:45161 IMAGE:5189970, complete cds.

BC033680
BC033680.1 GI:21707414

MGC

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 1624)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Sapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schneitz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, K.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Faney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (36), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 1624)

Strausberg, R.

Direct Submission

Submitted (02-JUL-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgaps@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland

Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@nsl.nih.gov

Akhtar, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,

Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Thurgood, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAX Plate: 68 Row: m Column: 1

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 6912327.

Location/Qualifiers

1. 1624

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

FEATURES

source

```

/clone="MGC:45161 IMAGE:5189970"
/tissue_types="Colon, Kidney, Stomach, adult, whole pooled"
/clone_lib="NIH_MGC_116"
/lab_host="DH10B"
/note="vector: pCMV-SORT6"
1..1624
/gene="DDAH1"
/db_xref="LocusID:23576"
/db_xref="LOCUSID:23576"
/db_xref="MIM:604743"
100..957
/codon_start=1
/product="DDAH1 protein"
/protein_id="AAH33680.1"
/db_xref="GI:21707415"
/db_xref="LocusID:23576"
/translation="MAGLHPAAGFGRATHAVRALPESLGHQALRSKAGEEVDVARAE
RQQLYVGLGSLQVLPADSLPDCVFEVDVAVCEETALITRCPASRKEV
DMKKEALEKQLATVEMKDNLTGTREFFVGLSKETNORGAELIADTFKD
YAVSTVPADGLHLKSPCSMAGPNLJAIGSSESAOKALXIMQMSDRHDKLITVDDI
ANCIYLINPNKSHVLLHRTPEYPESAKYIEKLDHMLIPVSMSELEKVDGLTCCS
VLINKKVD"
misc_feature
133..936
/note="Aminotransf, Region: Amidinotransferase. This
family contains glycine (EC:2.1.4.1) and inosamine
(EC:2.1.4.2) amidinotransferases, enzymes involved in
creatine and streptomycin biosynthesis respectively. This
family also includes arginine deiminases, EC:3.5.3.6.
These enzymes catalyse the reaction: arginine + H2O <=>
citrulline + NH3. Also found in this family is the
Streptococcus anti tumor glycoprotein"
/db_xref="CDD:pfam02274"

ORIGIN
Alignment Scores:
Pred. No.: 3,62e-122 Length: 1624
Score: 1439.00 Matches: 283
Percent Similarity: 99.65% Conservative: 1
Best Local Similarity: 99.30% Mismatches: 1
Query Match: 98.97% Indels: 0
DB: 9 Gaps: 0

US-09-889-733B-2 (1-285) x BC033680 (1-1624)
Qy 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
Db 100 ATGCGCGGGTCGGCCACCACCGCGCTCGCGCGGCCACCACCGCGCTGGCGGGCG 159
Qy 21 LeuProGluSerLeuGlyHisAlaLeuArgSerAlaGlyGlyGluGluValVal 40
Db 160 CTACCGGAGTCGTCGGCCAGCAGCGCTGAGAGGCCCAAGGCGGAGAGTGGACGTC 219
Qy 41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
Db 220 GCCCGCGGACCGGACGACCGAGCTCTAGCTGGGCGTGTGGGACGAGCTGGGGCTG 279
Qy 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
Db 280 CAGGTGTGTGAGTGTGGCGCGGACGAGCTTCGCGACTGCGCTTCGTGGGAGGACGTG 339
Qy 81 AlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgGlyLys 100
Db 340 GCCGTGTGTGCGGAGAGACGCGCCCTCATCCCGACCGCGCGCGCGGAGGAG 399
Qy 101 GlnValAspMetMetLysGluAlaLeuGluGlyLeuGlnLeuAsnLeuValGluMetLys 120
Db 400 GAGGTTGACATGATGAAGAAGCATTTAGAAAACTTCAGCTCAATATAGATGAGTAAA 459
Qy 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal 140
Db 460 GATGAAAAATCAACTTTAGATGCGGAGATGTTTATTACAGGCGAGAAATTTTGTG 519
Qy 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuAlaAspThrPheLysAsp 160

```

```

Db 520 GGCCTTTCCAAAAGGACAAATCAACAGGTCGTGAATCTTGGCTGATCTTTAAGGAC 579
Qy 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
Db 580 TATGCGTCTCCACAGTCCAGTGGCAGATGGTGTGCAATTTGAAGAGTTTTCGACGATG 639
Qy 181 AlaGlyProAsnLeuLeuAlaLeuGlySerSerGluSerAlaGlnLysAlaLeuLysile 200
Db 640 GCTGGGCTCAACCTGATGCAATTTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATC 699
Qy 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
Db 700 ATGCAACAGATGAGTGACCCACCGCTACGACAAATCACTGTGCTGATGATGACATGAC 759
Qy 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
Db 760 AACTGTATATATCTAATATATCCCAACAAAGGCGACGCTTGTGTCACCGAACCCCGGAA 819
Qy 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal 260
Db 820 GAGTATCCAGAAAGTGAAGGTTTATGAGAAATCAAGGACCATATGCTGCTGCTCCCTG 879
Qy 261 SerMetSerGluLeuGluValAlaAspGlyLeuLeuThrCysCysSerValLeuLeuLeu 280
Db 880 AGCATGCTGAACTGGAAAGTGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 939
Qy 281 LysLysValAspSer 285
Db 940 AAGAAAGTAGACTCC 954

RESULT 4
BD094073 1633 bp DNA linear PAT 27-AUG-2002
LOCUS Shear stress-responsive DNAs.
DEFINITION BD094073
ACCESSION BD094073
VERSION BD094073.1 GI:22639661
KEYWORDS WO 0125427-A/34.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Nojima, H., Yoshisue, H., Obayashi, M., Ota, T., Kawabata, A.,
Sakurada, K., Kuga, T., Sekine, S., Nakamura, Y. and Sugan, S.
Shear stress-responsive DNAs
Patent: WO 0125427-A 34 12-APR-2001.
KYOWA HAKKO KOGIO CO LTD, HIROSHI NOJIMA, HAJIME YOSHISUE, MASAYA
OBAYASHI, TOSHIO OTA, AYAKO KAWABATA, KAZUHIRO SAKURADA, TETSURO KUGA,
SUSUMU SEKINE, YUSUKE NAKAMURA, SUMIO SUGANO
OS Homo sapiens (human)
PN WO 0125427-A/34
PD 12-APR-2001
PF 02-OCT-2000 WO 2000JP006840
PR 01-OCT-1999 JP 99P 280976
PI HIROSHI NOJIMA, HAJIME YOSHISUE, MASAYA OBAYASHI, TOSHIO OTA, PI
AYAKO KAWABATA,
PI KAZUHIRO SAKURADA, TETSURO KUGA, SUSUMU SEKINE, YUSUKE NAKAMURA,
SUMIO SUGANO
PC C12N15/12.C07K14/435.C07K16/18.C12P21/02.C12Q1/68.A61K38/00,
PC A61K39/395
PC A61K48/00.A61P9/10.G01N33/50.G01N33/53
CC
CDS
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..1633
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

Alignment Scores:

```



```
Qy 101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
Db 623 GAGGTTGACATGATCAAGAAGCAATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAA 692
Qy 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheVal 140
Db 683 GATGAAATGCAACTTTAGATGGCGAGAGATGTTTTATTCACAGCAGAGAAATTTTTGTG 742
Qy 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp 160
Db 743 GGCCTTTCCAAAGAGCAAAATCAACGAGGTGCTGAAATCTTGCTGATACTTTTAAGGAC 802
Qy 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
Db 803 TATCAGCTCTCCACAGTGCCAGTGCCAGATGGGTTCATTTGAGAGATTTCTGCAGCATG 862
Qy 181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200
Db 863 CTGGGCTTAACCTGATCGCAATTTGGGTCTAGTGAATCTGCACAGAGAGGCTTAAGATC 922
Qy 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
Db 923 ATGCAACAGATGAGTGACCACTAGCAGCAAACTCACTGTGCTGATGATGATGACAGCA 982
Qy 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
Db 983 AACGTATATATCTAAATATATCCCAACAAAGGCGACGCTCTTGCTGCACCGAACCCCGAA 1042
Qy 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal 260
Db 1043 GAGTATCCAGAAAGTGCAAGGTTTATGAGAAACTGAAGACCATATGCTGATCCCGTG 1102
Qy 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuLeuAsn 280
Db 1103 AGCATGCTGAACGTGAAAGGTGATGGGTGCTCACTGCTGCTCACTGCTCACTGCTCACT 1162
Qy 281 LysLysValAspSer 285
Db 1163 AAGAAAGTAGACTCC 1177

RESULT 5
AX548046
LOCUS
DEFINITION
Sequence 22 from Patent WO02066654.
ACCESSION
AX548046
VERSION
AX548046.1 GI:25813142
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Astromoff, A., Au-Young, J., Baughn, M.R., Ding, L., Duggan, B.M.,
Forsythe, I. J., Gietzen, K.J., Griffin, J.A., Lee, E.A., Lu, Y.,
Richardson, T.W., Ring, H.Z., Sanjanwala, M.M., Swarnakar, A.,
Wallia, N.K., Warren, B.A., Xu, Y., Yue, H. and Zebajadian, Y.
TITLE
Drug metabolizing enzymes
JOURNAL
Patent: WO 02066654-A 22 29-AUG-2002;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
source
1. 4208
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7671089CB1"

ORIGIN
Alignment Scores:
Pred. No.: 1,296-121 Length: 4208
Score: 1439.00 Matches: 283
Percent Similarity: 99.65% Conservative: 1
Best Local Similarity: 99.30% Mismatches: 1
```

```
Query Match: 98.97% Indels: 0
DB: 6 Gaps: 0
US-09-889-733B-2 (1-285) x AX548046 (1-4208)
Qy 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
Db 368 ATGGCGGGGCTCGGCACACCCCGCCCTTCGGCCGGGCCACCCACGCCGCTGGTGGCGG 427
Qy 21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluValAspVal 40
Db 428 CTACCCGAGTCTCGGCAGCAGCACCGCTGAGAGCGCAAGGGCGAGAGGTGACGCTC 487
Qy 41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
Db 488 GCCCGCGCGGAACGCGACGACCCAGCTCTACGTGGCGCTGCTGGGCGAGCAGAGCTGGGG 547
Qy 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
Db 548 CAGGTGGTGGAGCTCCCGCGCAGCAGAGCCTTCGGGACTGCGTCTCTGCTGAGAGCGTG 607
Qy 81 AlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
Db 608 GCCGTGGTGTGGAGGAGACGCGCCCTCATCCCCCGGGCGCCGCGCGAGGAGAG 667
Qy 101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
Db 668 GAGGTTCATCATGATGAAGAAGCAATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAA 727
Qy 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheVal 140
Db 728 GATGAAATGCAACTTTAGATGGCGAGATGTTTTATTTCACAGGCGAGAGAAATTTTTGTG 787
Qy 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp 160
Db 788 GGCCTTTCCAAAGAGCAAAATCAACAGAGTGCTGAAATCTTGCTGATACTTTTAAGGAC 847
Qy 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
Db 848 TATGCACTCTCCACAGTGCCAGTGCGATGCGTATTTGAAGAGTTCTTCGACAGCATG 907
Qy 181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200
Db 908 GTTGGGCTTAACCTGATCGCAATTTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATC 967
Qy 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
Db 968 ATGCAACAGATGAGTGACCCCGCTACGACAAACTCACTGTGCTGATGACATAGACGCA 1027
Qy 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
Db 1028 AACTGTATATATCTAAATATCCCCAACAAAGGCGACGCTTCTGCTGCACCGAACCCCGGAA 1087
Qy 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal 260
Db 1088 GAGTATCCAGAAAGTGCAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCCGTG 1147
Qy 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuLeuAsn 280
Db 1148 AGCATGCTGAACTGGAAAGGTGATGGGTGCTCACTGCTGCTCACTGCTCACTGCTCACT 1207
Qy 281 LysLysValAspSer 285
Db 1208 AAGAAAGTAGACTCC 1222

RESULT 7
BC034505
LOCUS
DEFINITION
Mus musculus dimethylarginine dimethylaminohydrolase 1, mRNA (cdna)
clone MGC:30267 IMAGE:5134831, complete cds.
ACCESSION
BC034505
VERSION
BC034505.1 GI:21961623
KEYWORDS
MGC.
```


SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1120)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haie, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Viallano, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A.C., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 MEDLINE 22388257
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 1120)
 AUTHORS Strausberg, R.
 JOURNAL Direct Submission
 TITLE Submitted (24-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: angbcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 41 Row: 0 Column: 4
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.
 FEATURES
 source
 1..1120
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="MGC:30267 IMAGE:5134831"
 /tissue_type="liver, normal. 5 month old male mouse."
 /clone_lib="NCI CGAP_Li9"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1..1120
 /gene="Ddahl"
 /note="synonym: 2510015N06Rik"
 /db_xref="LocusID:69219"
 /db_xref="MGI:1916469"
 122..979
 CDS
 misc_feature
 155..958
 /note="Amidinotransf; Region: Amidinotransferase. This family contains glycine (EC:2.1.4.1) and inosamine (EC:2.1.4.2) amidinotransferases, enzymes involved in creatine and streptomycin biosynthesis respectively. This family also includes arginine deiminases, EC:3.5.3.6. These enzymes catalyze the reaction: arginine + H2O <=> citrulline + NH3. Also found in this family is the Streptococcus anti tumour glycoprotein"
 /db_xref="CDD:pfam02274"
 ORIGIN
 Alignment Scores:
 Pred. No.: 8,32e-118 Length: 1120
 Score: 1389.00 Matches: 269
 Percent Similarity: 97.89% Conservative: 10
 Best Local Similarity: 94.39% Mismatches: 6
 Query Match: 95.53% Indels: 0
 Dbs: 10 Gaps: 0
 US-09-889-733B-2 (1-285) x BC034505 (1-1120)
 Qy 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValAlaArgAla 20
 Db 122 ATGGCGCGCTCGGCCACCCCTCGGCTTGGCGGGCCACCCACCGCGTGTGCGGGCT 181
 Qy 21 LeuProGluSerLeuGlyGlnHisAlaLeuArgSerAlaGlyGluGluValAlaAspVal 40
 Db 182 CCGCGCGAGTCCCTGTGCGCCCGCCAGCGCTCGAGCGCTCGCAGCGCGAGGAGTTC 241
 Qy 41 AlaArgAlaGluArgGlnHisGlnLeuTyrrValGlyValLeuGlySerLysLeuGlyLeu 60
 Db 242 GCTCGCGCGAGCGCGCAGCAGAGCTCTACGTGGCGTGTGGCGCAGCAAGCTGGGCGTG 301
 Qy 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
 Db 302 CAGGTGTGAGCTGCGCGCGCAGAGAGCTCGCGAGCTCGCGTGTGTGCGGAGGAGCTG 361
 Qy 81 AlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLys 100
 Db 362 GCGTGTGTGCGGAGAGAGCGCCCTCATCCCGCGCGCGCGCGCGCGCAGGAG 421
 Qy 101 GluValAspMetLysGluAlaLeuGluLysLeuGlnLeuAlaValGluMetLys 120
 Db 422 GAGGTTGCATGATGAAGAGGCTTTGGAAAACTTCAGCTCAACATAGTAGATGAAA 481
 Qy 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal 140
 Db 482 GATGAAATGCACTTTGGATGTGGGAGCTCTATTTCAGCGCAGAGATTTTGTG 541
 Qy 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuLeuAlaAspThrPheLysAsp 160
 Db 542 GGCCTTTCCAAAAGAACAAATCAACAGAGTGTGTAATCTTGGCTGATCTTTAAGGAC 601
 Qy 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
 Db 602 TAGCAGCTCTCTACAGTCCCTGTGGCGGATCTTTGCGATTTAAGATTTCTGCAGCATG 661
 Qy 181 AlaGlyProAsnLeuLeuAlaLeuGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200
 Db 662 GCGGAGCCCAACCTGATTGCAATAGGCTCCAGCGAATCTGCACAGAGGCGCCTCAAGATC 721

QY 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
Db 722 ATGCAACAGATGAGTGACCATCGTTATGACAAGCTCACTGTACCGACGACGATGGCGCC 781
QY 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
Db 782 AACTGCATATATCTAAATATCCCGCAGCAAGGCGATGCTTGCTGCACCGCAACCCAGAA 841
QY 241 GluTyrProGluSerAlaLysValTyrGlnLysLeuLysAspHisMetLeuIleProVal 260
Db 842 GAGTACCCAGAAAGCGCAAGAGCTCTATGAGAAATCAAGGACCATCTACTGATCCCTGTG 901
QY 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuLeuAsn 280
Db 902 AGCAACTCGGAGATGGAAGGTGAGGGCTTGCTCACTGCTGCTGCTGCTGCTGCTGCTG 961
QY 281 LysLysValAspSer 285
Db 962 AAGAAGATAGACTCC 976
RESULT 8
AX683109
LOCUS AX683109 3008 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 83 from Patent EP1279744.
ACCESSION AX683109
VERSION AX683109.1 GI:29370146
KEYWORDS Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1
AUTHORS Brooksbank, R.A., Dixon, A.K., Lee, K. and Pinnock, R.D.
TITLE Identification and use of molecules implicated in pain
JOURNAL Patent: EP 1279744-A 83 29-JAN-2003;
WARNER-LAMBERT COMPANY (US)
FEATURES
source
1. 3008
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
/note="N-G, N-G dimethylarginine dimethylaminohydrolase"
ORIGIN
Alignment Scores:
Pred. No.: 2.56e-116 Length: 3008
Score: 1379.00 Matches: 268
Percent Similarity: 96.84% Conservative: 8
Best Local Similarity: 94.04% Mismatches: 9
Query Match: 94.84% Indels: 0
Gaps: 6
US-09-889-733B-2 (1-285) x AX683109 (1-3008)
QY 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
Db 432 ATGGCCGGCCCTCAGCCACCCCTCGGTCTTCGGCCGGCCACCCACCGCGTGTGGCGGCT 491
QY 21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluGluValAspVal 40
Db 492 CGGCCGAGTCCCTGTGCGCGCACCGCTGAGCGCTCCCGAGCGGAGAGGTGATTC 551
QY 41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
Db 552 GCTCGCGTGTAGCGGCACACACAGCTCTACGTGGCGGTGCTGGCGAGCAAGCTGGGCGTG 611
QY 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
Db 612 CAGGTGGTGCAGCTGCGCGCCGACGAGAGCTGCTGACTGCTGCTGCTGCTGCTGCTGCTG 671
QY 81 AlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100

Db 672 GCCGTGGTGTGCGAGGAGACCGCCCTCATCACCCCGCGGCGCCCTAGCGCGAGGAAG 731
QY 101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
Db 732 GAGGTGTGATCATGATCAAGAGCGCTTTGGAAAACCTTCAGCTCAACATAGTAGATGAAA 791
QY 121 AspGluAsnAlaThrLeuAspGlyClyAspValLeuPheThrGlyArgGluPhePheVal 140
Db 792 GATGAAAATGCAACCTTAGATGGTGGGACGCTCTATTTCACAGGCGAGAGTTTGTG 851
QY 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp 160
Db 852 GGCCTTCCAAAAGACAAATCAACGAGGTGCTGAGATCTTGGCTGATACTTTCAAGGAC 911
QY 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
Db 912 TAGCGAGTTTCACAGTCCCGTGGCCGATCTTTTTCATTTAAAGAGTTTTCGACGATG 971
QY 181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200
Db 972 GTGGCCCACTGATCGCAATAGGCTCCAGTGATCTGCGCAGAGGCCCTCAAGATC 1031
QY 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
Db 1032 ATGCAACAGATGAGTGACCGCCGTTATGACAGCTCACTGTACCGGACGACATGCCGCC 1091
QY 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
Db 1092 AACTGTATATATTTAAATATCCCGAAGGCGATGCTTGTGTCACCGAACCACCA 1151
QY 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal 260
Db 1152 GAGTACCCAGAAAGCGCAAGGTTTATGAGAAGCTCAAGGACCATCTACTATCCCTGTG 1211
QY 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuIleAsn 280
Db 1212 AGCAATTCGAGATGGAAGGTGAGCGCTTGTCTGCTGCGACCGAACCACCA 1271
QY 281 LysLysValAspSer 285
Db 1272 AAGAAGACAGACTCT 1286
RESULT 9
D86041
LOCUS D86041 3008 bp mRNA linear ROD 06-FEB-1999
DEFINITION Rattus norvegicus mRNA for N-G,N-G-dimethylarginine dimethylaminohydrolase, complete cds.
ACCESSION D86041
VERSION D86041.1 GI:1906799
KEYWORDS N-G,N-G-dimethylarginine dimethylaminohydrolase.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (sites)
AUTHORS Kimoto, M., Sasakawa, T., Tsuji, H., Miyatake, S., Oka, T., Nio, N. and Ogawa, T.
TITLE Cloning and sequencing of cDNA encoding NG,NG-dimethylarginine dimethylaminohydrolase from rat kidney
JOURNAL Biochim. Biophys. Acta 1337 (1), 6-10 (1997)
MEDLINE 97157050
PUBMED 9003431
REFERENCE 2 (bases 1 to 3008)
AUTHORS Kimoto, M.
TITLE Cloning and sequencing of cDNA encoding G-N, G-N-dimethylarginine dimethylaminohydrolase from rat kidney
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 3008)
AUTHORS Kimoto, M.
TITLE Direct Submission
JOURNAL Submitted (14-JUN-1996) Masumi Kimoto, The University of Tokushima, School of Medicine, Nutrition; Kuramoto 3, Tokushima, Tokushima

770, Japan (E-mail: kimoto@nutr.med.tokushima-u.ac.jp,
Tel.0896-33-7088, Fax:0896-33-7089)

FEATURES

source

1. 3008
/location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Wistar"
/db_xref="taxon:10116"
/sex="female"
/tissue_type="kidney"
432. 1289
/note="DDAH"
/codon_start=1
/product="N-G,N-G-dimethylarginine dimethylaminohydrolase"
/protein_id="BAAL993.1"
/db_xref="GI:1906300"
/translation="MAGLSHVSFCRATHAVRAPPELSCRHALRRSQGEEVDFARAE
ROHLYVGLSKGLQVQVPADESLDFVAVVCEETALIIRPGAPSRKKEV
DMKEALEKQINIVEMDENATDGGDLFTGREFFVGLSKTRNQRGALEILATFD
YAVSTVPADSLHLKSPGAPNLIAIGSSSAQKALKINQMSDHRDKLITVPDDM
AANCYVLPKSGHVLHRTPEVPESAKVYEKLDHLLIPVNSEMEKVDGLLTCCS
VFINKTDS"

ORIGIN

Alignment Scores:
Pred. No.: 2,56e-116 Length: 3008
Score: 1379.00 Matches: 268
Percent Similarity: 96.84% Conservative: 8
Best Local Similarity: 94.04% Mismatches: 9
Query Match: 94.84% Indels: 0
DB: 10 Gaps: 0

US-09-889-733B-2 (1-285) x D86041 (1-3008)

Qy 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
Db 432 ATGGCGCGGCTCAGCACCCCTCGTCTTCGGCGGGGACCCACCGCGTGTGGCGGCT 491
Qy 21 LeuProGluSerLeuGlyHisAlaLeuArgSerAlaGlyGlyGluValVal 40
Db 492 CGCGCGAGTCTGTCGCGCACCGCTGAGCGCTCCACGGGCGAAGAGTGGATTTC 551
Qy 41 AlaArgAlaGluArgGlnHisGlnLeuThrValGlyValLeuGlySerLeuGlyLeu 60
Db 552 GTCGCGCTGAGCGGCAGCACCGCTCTACGTGGCGGCTGCTGGGCGAAGAGTGGCGGCTG 611
Qy 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
Db 612 CAGGTGTGACGTCCGCGCGACGAGAGCGCTGCTGACTGGGTTCGTGGAGGACGTG 671
Qy 81 AlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLys 100
Db 672 GCGGTGGTGTGGAGGAGACCGCTCTATCCCGCGCGGCGGCGCTAGCGCGAGGAAG 731
Qy 101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnLeuValGluMetLys 120
Db 732 GAGGTTCATGATGAAGAGCGCTTGGAAAACTTCAGCTCAACATAGTAGAGATGAA 791
Qy 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheVal 140
Db 792 GATGAAAATGCAACCTTAGTGTGGGACGCTCTATTACAGGCGAGAGATTTTGTG 851
Qy 141 GlyLeuSerLysArgThrAsnGlnArgGlyValGluLeuAlaAspThrPheLysAsp 160
Db 852 GGCCTTTCMAAGGACAAATCAACAGGTGCTGAGATCTTGGCTGATCTTTCAGGAC 911
Qy 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
Db 912 TACGCAGATTTCACAGTCCCGGTGGCGGATCTTTCATTTAAAGAGATTTTCGACGATG 971
Qy 181 AlaGlyProAsnLeuLeuAlaLeuGlySerGluSerAlaGlnLysAlaLeuLysLe 200
Db 972 GCTGCGCCCACTGATCGCAATAGGGTCCAGTGAATCTGCGCAGAGGCGCCCTCAAGATC 1031

Qy 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
Db 1032 ATGCAACAGATGAGTGACCACTTATGCAAGCTCCTGTACCGACGACATGGCGGCC 1091
Qy 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
Db 1092 AACTGTATATTAATATATCCCAAGCAAGGCGATGCTTGTGTCACCGAAGCCACGAA 1151
Qy 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuLeuProVal 260
Db 1152 GAGTACCCAGAAAGCGCAAGGTTTATGAGAACTCAAGGACCATCTACTGTATCCCTGTG 1211
Qy 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysSerValLeuLeuAsn 280
Db 1212 AGCAATCTGAGATGAAGAGGTGGCGGCTTCTCCTGCTGCTCCGTTTATTATAC 1271
Qy 281 LysLysValAspSer 285
Db 1272 AAGAAGACAGACTCT 1286

RESULT 10

BC043235

LOCUS

DEFINITION

BC043235

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 3737)

Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L.,

Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S.,

Carninci P., Prange C., Raja S.S., Loquellano N.A., Peters G.J.,

Abramson R.D., Mullaby S.J., Bosak S.A., McEwan P.J.,

McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,

Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

Fahney J., Helton E., Kettman M., Madan A., Rodriguez S.,

Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,

Boffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,

Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

Butterfield Y.S., Krzywinski M.I., Skalska U., Smalusz D.E.,

Schnerch A., Schein J.E., Jones S.J. and Marra M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL

MEDLINE

PUBMED

REFERENCE

2 (bases 1 to 3737)

Strausberg R.

Direct Submission

Submitted (09-JAN-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.ncbi.nlm.nih.gov>

On Sep 16, 2003 this sequence version replaced gi:28175756.

Contact: MGC help desk

Email: cgabs@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)

CDNA Library arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre.

REMARK

COMMENT

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Kryzyski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lilias Prabhu, Parvaneh Saedi, Jacqueline
Schein, Duane Snailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 76 Row: j Column: 15.

FEATURES

source

Location/Qualifiers
1..3737
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5295723"
/issue_type="Testis"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="vector: plusescript"
<1..3737
/gene="DDAH1"
/note="synonym: DDAH"
/db_xref="LocusID:23576"
/db_xref="MIM:604743"
<1..741
/gene="DDAH1"
/codon_start=1
/product="DDAH1 protein"
/protein_id="AAH43235.2"
/db_xref="GI:34783629"
/db_xref="LocusID:23576"
/db_xref="MIM:604743"
/translation="VARAERQHQLYVGLSKGLQVLPAPESLPDCCVVEDVAV
CEETALTTPGAPRRKEVKEALEKQINIVEMKDNATLDGGDVLFTGREFVVG
LSKRTNQRGAIEHLTKFVAVSTVPVADGLHLKSCFSGMAGPNLIIAGSSSAQKALK
IPQMSDRYDKLTPVDDIANCILYLNPNKGVHLLHRTPEYPESAKVYKELKDEL
IPVMSLEKVDGLTCCSVLNKKVDS"
1..720
/gene="DDAH1"
/note="Aminotransf; Region: Aminotransferase. This
family contains glycine (EC:2.1.4.1) and inosamine
(EC:2.1.4.2) aminotransferases, enzymes involved in
creatine and streptomycin biosynthesis respectively. This
family also includes arginine deiminases, EC:3.5.3.6.
These enzymes catalyse the reaction: arginine + H2O <=>
citrulline + NH3. Also found in this family is the
Streptococcus anti tumor glycoprotein"
/db_xref="CDD:pfam02274"

gene

CDS

misc_feature

ORIGIN

Alignment Scores:
Pred. No.: 1,17e-104 Length: 3737
Score: 1253.00 Matches: 246
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.18% Indels: 0
DB: 9 Gaps: 0

US-09-889-733B-2 (1-285) x BC043235 (1-3737)

Qy 40 VallalaArgAlaGluArgGlnLeuTyValGlyValLeuGlySerLysLeuGly 59
Db 1 GTGCGCCCGCGGACCGGACGACGCTTACGCTGGCGCTGCTGGCGACGACGCTGGGG 60
Qy 60 LeuGlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAsp 79
Db 61 CTGAGGTTGGAGCTCTCCGCGCGACGAGAGCTTTCGGACTCGCTTCGTTGGAGGAC 120

Qy 80 VallalaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArg 99
Db 121 GGGCGCGTGGTGTGGAGAGAGAGCGCCCTCATCCCCACCGCGCGCGCGCGGAGG 180
Qy 100 LysGluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMet 119
Db 181 AAGGAGGTTCATCATGATGAAGAAGCATTAGAAAAACCTCAGCTCAATATAGTAGAGATG 240
Qy 120 LysAspGluAsnAlaThrLeuAspGlyClyAspValLeuPheThrGlyArgGluPhePhe 139
Db 241 AAAGATGAATATGCAACTTTAGATGCGGAGATGTTTATTTCACAGGCGAGAGAAATTTT 300
Qy 140 ValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLys 159
Db 301 GTGGGCGCTTCCAAAAGGACCAATCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAG 360
Qy 160 AspTyAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSer 179
Db 361 GACTATGCAGTCTCCACAGTGCAGTGCAGATGGTGGTTCATTTTGAAGAGTTTTCGACG 420
Qy 180 MetAlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLys 199
Db 421 ATGGTGGGCGCTTACCTGATCGCAATTTGGTCTAGTGAATCTGCACAGAGGCCCTTAAG 480
Qy 200 IleMetGlnGlnMetSerAspHisArgTyArgLysLeuThrValProAspAspIleAla 219
Db 481 ATCATGCAACAGATGAGTGCACCGCTACGACCAAACTCACTGTGCTGATGACATAGCA 540
Qy 220 AlaAsnCysIleTyLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrPro 239
Db 541 GCAAACTGTATATATCTAATATCCCAACAAAGGCGACGCTTCTGCTGACGACGACCCG 600
Qy 240 GluGluTyProGluSerAlaLysValTyGluLysLeuLysAspHisMetLeuIlePro 259
Db 601 GAAGAGTATCCAGAAAGTCAAAAGGCTTTATGAGAACTGAAGACCATATGCTGATCCCC 660
Qy 260 ValSerMetSerGluLeuGluLysValAspGlyLeuThrCysCysSerValLeuIle 279
Db 661 GTGAGCATGCTGAACTGGAAAGGTGATGGCTGCTGCCTGCTGCTGCTGCTGCTGCTGCT 720
Qy 280 AsnLysLysValAspSer 285
Db 721 AACAAAGAAAGTAGACTCC 738

RESULT 11
AB095027
LOCUS
DEFINITION
AB095027
AB095027.2 GI:32306458
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

AB095027
Gallus gallus mRNA for dimethylarginine dimethylaminohydrolase I,
complete cds.
AB095027.2 GI:32306458
Gallus gallus (chicken)
Gallus gallus
Archaeopteryx; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Phasianinae; Gallus.
Mishima, T., Hamada, T., Ui-Tei, K., Takahashi, F., Miyata, Y.,
Imaki, J., Suzuki, H. and Yamashita, K.
The expression of DDAH I in the chick and rat embryo
Unpublished
2 (bases 1 to 1150)
Mishima, T., Hamada, T., Ui-Tei, K., Takahashi, F., Miyata, Y.,
Imaki, J., Suzuki, H. and Yamashita, K.
Direct Submission
Submitted (30-OCT-2002) Takuya Mishima, Nippon Medical School,
Department of Anatomy; Sendagi 1-1-5, Bunkyo, Tokyo 113-8602, Japan
(E-mail: tmishima@nms.ac.jp, Tel: 81-3-3822-2131 (ex.5315))
On Jun 26, 2003 this sequence version replaced gi:30268647.
Location/Qualifiers
1..1150
/organism="Gallus gallus"

```

/mol_type="mRNA"
/db_xref="taxon:9031"
/tissue_type="embryos yolk sac"
/clone_lib="Chick Yolk Sac Lambda cDNA Library (Stratagene #937406)"
/dev_stages="Five-days-old chick"
40..903
CDS
/codon_start=1
/product="dimethylarginine dimethylaminohydrolase I"
/protein_id="BAC76002.1"
/db_xref="GI:30268648"
/translation="WAGLGGGPAFGRCTHVAVRALPSLCRQALRSTAGPEVDFAA
EREHLYVGRGLQVLPADESLPCFVEDAAVCEESTALLTRPGASRRKE
VEAMKRVLSLNANVMVDNATLDGDLVLTGREFFVGLSKRTNQGAELILTFP
KDYAVTVPVDSHLKFKSCWAGNPLIAGSSEAOKALKTMOQMSDHRDKLITVPD
DAAANCIYLNIPSKHVLHRAPEYPSAKVFEKLKDHLIPIANTELEKVDGALT
CSVLINKTSEL"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 2e-104 Length: 1150
Score: 1243.00 Matches: 245
Percent Similarity: 91.17% Conservative: 13
Best Local Similarity: 86.57% Mismatches: 23
Query Match: 85.49% Indels: 2
DB: 5 Gaps: 2
US-09-889-733B-2 (1-285) x AB095027 (1-1150)
QY 1 MetAlaGlyLeu--GlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArg 19
Db 40 ATGCGCGGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 99
QY 20 AlaLeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluValAlaAsp 39
Db 100 GCGCTGCCGAGTGGCTGTGCGCGCGAGGCGTGTGCGCAGCAGCGCGCGCGCGCGAG 159
QY 40 ValAlaArgAlaGluArgGlnHisGlnLeuTyValGlyValLeuGlySerIysLeuGly 59
Db 160 TTGCTCGCGCGGAGGCGGAGCATGAGTGTACGTGGCGGTCTACGGGGCAAGCTGGGG 219
QY 60 LeuGlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAsp 79
Db 220 CTGCAGTCTGGAGTGGCGGCTGACGAGAGCCCTCCGAGCTGCTTCTGCGAGGAGC 279
QY 80 ValAlaValValCysGluGluThrAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 99
Db 280 GCGCGCGTGTCTGCGAGGAGAGCGCGCTGTCTCACCGCGCGCGCGCGCGCGCGCGAG 339
QY 100 LysGluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnLeuValGluMet 119
Db 340 AAGAGTGTGAACCATGAAGAGAGTACTAGAAAGCCTTAACCTCAACATGATGAGATG 399
QY 120 LysAspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGlu---Phe 138
Db 400 GTAGATGAAATCAACCTTGGATGGTGAGAGCTCTTATTTACAGCGCAGAGAAATTTT 459
QY 139 PheValGlyLeuSerIysArgThrAsnGlnArgGlyAlaGluLeuLeuAlaAspThrPhe 158
Db 460 TTTGTAGGCTTTCCAAAGCGGCAAAATCAACGTTGGGCGAGAAATCTGGCTGACACAT 519
QY 159 LysAspTyAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCys 178
Db 520 AAGATATATGCTCTCCACAGTCCCTGTTTCATGATCTTTGATCTGAAGAGTTTTCG 579
QY 179 SerMetAlaGlyProAsnLeuLeuAlaLeuGlySerSerGluSerAlaGlnLysAlaLeu 198
Db 580 AGCATGGCTGGACAAACCTGATTTGTTGCTGAGTCAAGTGAAGTGAAGTGAAGTGAAG 639
QY 199 LysIleMetGlnGlnMetSerAspHisArgTyValAspLysLeuThrValProAspPhe 218
Db 640 AAGACCATGCAAGATGAGCGACCCAGCTACCAAGCTGACGTGCGTGCACGATGCC 699

```

```

QY 219 AlaAlaAsnCysIleTyValLeuAsnLeuProAsnLysGlyHisValLeuLeuHisArgThr 238
Db 700 GCGCAAACTGATCTACTTAAACATCCAGCAAGAGCGACGCTCTGCTGCGACCGGCC 759
QY 239 ProGluGluTyProGluSerAlaLysValTyGluLysLeuLysAspHisMetLeuLeu 258
Db 760 CTTGAGAGTATCCAGAGAGCGCAAGAGTTTTCGAAAAATCGAAGGACCACTGCTGATC 819
QY 259 ProValSerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeu 278
Db 820 CCATAGCCCAACACAGACTGGAGAAAGTAGATGGGCGCACTCACTGTTGCTGTGCTT 879
QY 279 IleAsnLys 281
Db 880 ATTAACAAA 888
RESULT 12
LOCUS BC056074 1344 bp mRNA linear VRT 04-NOV-2003
DEFINITION Xenopus laevis dimethylarginine dimethylaminohydrolase 1, mRNA
(GDNA clone MGC:69055 IMAGE:4965611), complete cds.
ACCESSION BC056074
VERSION BC056074.1 GI:33417137
KEYWORDS MGC.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 1344)
AUTHORS Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
and Richardson,P.
TITLE Genetic and genomic tools for Xenopus research: The NIH Xenopus
Initiative
JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
MEDLINE 22341132
PUBMED 12454917
REFERENCE 2 (bases 1 to 1344)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,P.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
PUBMED 12477932
REFERENCE 3 (bases 1 to 1344)
AUTHORS Klein,S. and Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-2003) National Institutes of Health, Xenopus Gene
Collection (XGC), National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
NIH-MGC Project
Contact: XGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Igor Dawid
cDNA Library Preparation: Life Technologies, Inc.
REMARK
COMMENT

```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNML)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNML at: <http://image.llnl.gov>
 Series: FRAX Plate: 129 Row: J Column: 20
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, Similarity but not identity to protein.

FEATURES

Location/Qualifiers
 1..1344
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="MGC:69055 IMAGE:4965611"
 /tissue type="Spleen, adult Xenopus"
 /clone_lib="NICHD XGC_Spl"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 1..1344
 /gene="ddahl-prov"
 /note="synonym: MGC69055"
 /db_xref="LocusID:380577"
 84..938
 /codon_start=1
 /product="ddahl-prov protein"
 /protein_id="AAH56074.1"
 /db_xref="GI:33417138"
 /db_xref="LocusID:380577"
 /translation="MAGLALSGFGRCTHVIKRAIPDLSLQALRIKGSVVDISLARK
 QHEQVEVLKGLSVTELPADSLDPFVDVAVCEETALITRPGAPSRKETE
 AVKALQDLQNVEMTDENALDGGDLVFTGREFVGLSKRTNQRAEILADTFKDY
 AVSTVPVETLHLKSFCSMAGPNLIAIGSSEAOKALKTWQMSDRHYDKLTLPDAA
 ANCYLAIKPSKGVHLLHAPDFPESAKVEKLEHMLFPLNTELAKVDCGLTCCSI
 LINKKSL"

114..917
 /note="Aminotransf; Region: Amidinotransferase. This
 family contains glycine (EC:2.1.4.1) and inosamine
 (EC:2.1.4.2) amidinotransferases, enzymes involved in
 creatine and streptomycin biosynthesis respectively. This
 family also includes arginine deiminases, EC:3.5.3.6.
 These enzymes catalyse the reaction: arginine + H2O <=>
 citrulline + NH3. Also found in this family is the
 Streptococcus anti tumour glycoprotein"
 /db_xref="CDD:pfam02274"

misc_feature

ORIGIN

Alignment Scores:
 Pred. No.: 8,96e-96 Length: 1344
 Score: 1149.50 Matches: 225
 Percent Similarity: 86.85% Conservative: 26
 Best Local Similarity: 77.85% Mismatches: 33
 Query Match: 79.06% Indels: 5
 DB: 5 Gaps: 2

US-09-889-733B-2 (1-285) x BC056074 (1-1344)

QY 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValAlaValAla 20
 Db 84 ATGGCTGACCTAGCT---CTGTCAGGCTTTGGGAGATGCACCCAGTTATAGTTCGTGCA 140
 QY 21 LeuProGluSerLeuGlyGlnHisAlaLeuArgSerAlaLysGlyGluGluValAspVal 40
 Db 141 ATTCTCTGACTCTTTATGTCAAGAAGCTCTGAGGATAGAGAAGGCAAGTGGTGGATATC 200
 QY 41 AlaArgAlaGluGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
 Db 201 TCCCTGGCTAGGAACAACATGAGCAGTATGTTGAGGTGCTGAAGAATAAACTGGGCGCTG 260

QY 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
 Db 261 ACTGTGATAGACTTCCAGCAGATGAATCCCTCCAGATGGCCCTTTGTGGAAGATGTG 320
 QY 81 AlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLys 100
 Db 321 GCTGTGTGTGTGTGAGAACTGCCCTGATTACCAGACCTGGGGACCTAGCAGGAGGAAA 380
 QY 101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
 Db 381 GAGACTGAAGCTGTGAGAAGCACTGCAGGATCTTCAGCTAAATGTGGTTGAATGACA 440
 QY 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal 140
 Db 441 GATGAAATGCTACATTAGATGAGAGAGATGTTTGTTTTACAGGAGGGAATCTTTGTG 500
 QY 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp 160
 Db 501 GUTTTGTGAAAGGACCAATCAAGTGTGTGCTGAAATCTGGCAGATACCTTTAAGGAT 560
 QY 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
 Db 561 TATGCAGTTTCCACTGTGCCAGTGTGCGAGACATTCACATTTGAAGAGTTTCTGCAGCATG 620
 QY 181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200
 Db 621 GCTGGACCAAACTCATTCGCCATTCGCTGAGTGAAGCTGCCCAAAAGCTCTAAAGACC 680
 QY 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaIle 220
 Db 681 ATGCAGCAGATGAGTGTACCGCTACGACAGCTGACATTCCTGCTGATAATGCTGCAGCT 740
 QY 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
 Db 741 AACTGCATTTATTTAAACATTCCAAGCAAGGCGCATGTTTACTCATCTTCCACTGAA 800
 QY 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal 260
 Db 801 GATTTTCCAGAGTGCAGAAAGTGTGTAAGAACTAAAGAACACATGCTGTTCCACTT 860
 QY 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuLeuAsn 280
 Db 861 CCTAACACCCGAGCTTGCAAAAGTTGATGATGCTGTGACCTGCTGCTCATCTTATTAAAC 920
 QY 281 LysLys-----ValAspSer 285
 Db 921 AAAAAAAGTCTCTGAGTTGATTCA 947

RESULT 13

HSM808292 3908 bp mRNA linear PRI 30-AUG-2003
 LOCUS Homo sapiens mRNA, CDNA DKFZp686N2176 (from clone DKFZp686N2176).

DEFINITION EX648145

ACCESSION BX648145.1 GI:34367304

VERSION Homo sapiens (human)

KEYWORDS Homo sapiens

SOURCE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 3908)

Ausio, J., Krieger, S., Regier, T., Rittmüller, C., Schwager, B.,

Newes, H.W., Well, B., Amid, C., Osanger, A., Fob, G., Han, M. and

Wiemann, S.

The German Human cDNA Consortium

Direct Submission

Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764

Neuberberg, GERMANY

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKPZ); Email s.wiemann@dkfz-heidelberg.de;

Heidelberg/Germany) within the cDNA sequencing consortium of the

German Genome Project.

COMMENT

This clone (DKFZp686N2176) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clonerzpd.de Further
information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

source
1. 3908
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686N2176"
/tissue="human endometrium carcinoma cell line"
/clone_lib="686 (synonym: hlec3). Vector pSporti_SfiI; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"

ORIGIN

Alignment Scores:
Pred. No.: 4.57e-76 Length: 3908
Score: 941.00 Matches: 184
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 64.72% Indels: 0
DB: 9 Gaps: 0

US-09-889-733B-2 (1-285) x HSM808292 (1-3908)

Qy 102 ValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAsp 121
Db 384 GTTCACATGATGAAGAAGCATATAGAAAACCTTCAGCTCATATAGTAGAGATGAAGAT 443
Qy 122 GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheValGly 141
Db 444 GAAATGCACTTTAGATGGCGAGATGTTTATTACAGCGCAGAGAATTTTGTGGGC 503
Qy 142 LeuSerLysArgThrAsnGlnArgGlyValGluLeuAlaAspThrPheLysAspTyr 161
Db 504 CTTTCCAAAAGGACAAATCAACAGGGTGCTGAAATCTTGCTGTATCTTTTAGGACTAT 563
Qy 162 AlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAla 181
Db 564 GCAGTCTCCACAGTGCCAGTGCGCAGATGGGTGCATTTGAAGAGTTTCTGCAGCATGGCT 623
Qy 182 GlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMet 201
Db 624 GGGCTTAACCTGATCGCAATGGGTCTAGTGAATCTGCACAGAGGCCCTTAGATCATG 683
Qy 202 GlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsn 221
Db 684 CAACAGATGATGACCCCGCTACGACAACTCACTGTGCTGTATGATGACAGCAAAAC 743
Qy 222 CysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGluGlu 241
Db 744 TGTATATATCTAAATATATCCCAAAAGGGCAGCTCTGTGACCCGAAACCCCGGAAGAG 803
Qy 242 TyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProValSer 261
Db 804 TATCCAGAAATGCAAGGTTTATGAGAACTGAGGACCATATGCTGATCCCCGTGAGC 863
Qy 262 MetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuIleAsnLys 281
Db 864 ATGCTGAATCTGAAAGGTGGATGGGTGCTACCTGCTGCTCAGTCTTAATTAACAAG 923
Qy 282 LysValAspSer 285
Db 924 AAGTAGACTCC 935

RESULT 14

AK128909 2289 bp mRNA linear ROD 09-SEP-2003
LOCUS
DEFINITION Mus musculus cDNA fis, clone TRACH2023479, highly similar to
NG-NG-DIMETHYLARGININE DIMETHYLAMINOHYDROLASE 1 (EC 3.5.3.18).
ACCESSION AK128909

VERSION

AK128909.1 GI:34536566

KEYWORDS

oligo capping; fis (full insert sequence).

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,
Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A.,
Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K.,
Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y.,
Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K.,
Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.

TITLE

NEDO cDNA sequencing project

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 2289)

AUTHORS

Isogai, T. and Yamamoto, J.

TITLE

Submitted (15-JUL-2003) Takao Isogai, FUJ Project (HRI Team); 2-6-7

JOURNAL

Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO cDNA sequencing project supported by Ministry of Economy,
Trade and Industry of Japan; cDNA full insert sequencing; Research
Association for Biotechnology (RAB); cDNA library construction;
Helix Research Institute (HRI) (supported by Japan Key Technology
Center etc.); 5'- & 3'-end one pass sequencing; RAB, HRI, and
Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing; HRI and
RAB; annotation: HRI and RAB.

FEATURES

source

1. 2289
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="TRACH2023479"
/note="Cloning vector: pME18SFL3"

ORIGIN

Alignment Scores:

Pred. No.: 1.54e-73 Length: 2289
Score: 910.00 Matches: 176
Percent Similarity: 98.3% Conservative: 5
Best Local Similarity: 95.65% Mismatches: 3
Query Match: 62.59% Indels: 0
DB: 10 Gaps: 0

US-09-889-733B-2 (1-285) x AK128909 (1-2289)

Qy 102 ValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAsp 121
Db 29 GTTGACATGATGAAGAAGCCTTTGGAATAACTTCAGCTCAACATAGTAGAGATGAAGAT 88
Qy 122 GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheValGly 141
Db 89 GAAATGCACTTTGGATGGTGGGACGCTCTATTCACAGGCAGAGATTTTGTGGGC 148
Qy 142 LeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyr 161
Db 149 CTTTCCAAAGAACAATAACACAGAGTGCTGAAATCTTGCTGTATCTTTTAGGACTAC 208
Qy 162 AlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAla 181
Db 209 GCAGTCTCTACAGTCCCTGTGGCCGATTTCTTGCATTTAAAGAGTTTCTGCAGCATGGCC 268
Qy 182 GlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMet 201
Db 269 GGACCCCAACCTGATTGCAATAGGGTCCAGCAATCTGCACAGAGCCCTCAAGATCATG 328
Qy 202 GlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsn 221

Db 329 CAACAGATGAGTACCATCTGTTATGACAAAGCTCACTGTACCCGACGACATGGCCGCCAAC 388
 QY 222 CysIIeTyLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGluGlu 241
 Db 389 TGCATATATCTTAATATCCCGACGAAAGGCGATGCTTGTCTGACCGAACCCGAGAG 448
 QY 242 TyrProGluSerAlaLysValTyrgLysLeuLysAspHisMetLeuLeuProValSer 261
 Db 449 TACCAGAAAGCGCAAGGCTCATGAGAAACTCAAGGACCATCTACTGATCCCTGTGAGC 508
 QY 262 MetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuLeuAsnLys 281
 Db 509 AACTCGGAGATGAAAGAGTGACGGCTGCTCACCTGCTGCTGCTTTTATTAAACAAG 568
 QY 282 LysValAspSer 285
 Db 569 AAGATAGACTCC 580
 RESULT 15
 BC060266 3420 bp mRNA linear ROD 25-NOV-2003
 LOCUS Mus musculus dimethylarginine dimethylaminohydrolase 1, mRNA (cDNA
 DEFINITION clone IMAGE:5698949), complete cds.
 ACCESSION BC060266
 VERSION
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 3420)
 AUTHORS Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zebberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, I., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Schaeetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Abramson, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Carmona, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Willeton, D.K., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Vortale, D.K., Murny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 TITLE
 JOURNAL
 MEDLINE
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 3420)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (20-OCT-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 cDNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo

Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandon, Anna-Liisa Prabh, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smallos, Michael Smith, Lorraine Spence, Jeff Scott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 132 Row: a Column: 19
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis
 This clone has the following problem: The cds is short compared to
 the longest cds in the locus.
 FEATURES
 source Location/Qualifiers
 1..3420
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAG:5698949"
 /tissue_type="Brain, whole, 18.5 dpc mouse embryo"
 /clone_lib="NIH_EMAP_F00"
 /lab_host="DH10B"
 /notes="Vector: pYX-ASC"
 1..3420
 /gene="Ddahl1"
 /notes="synonym: 2510015N06Rik"
 /db_xref="LocusID:69219"
 /db_xref="MGI:1516469"
 360..680
 /codon_start=1
 /product="Ddahl1 protein"
 /protein_id="AAH60266.1"
 /db_xref="GI:37805214"
 /db_xref="LocusID:69219"
 /translations="MAGPMLIAGSSSEKALKINQMDSREYDKLTPDDMAANCI
 YLNPISKGHVLLHRTPEEYPSAKYKULHLLIPVSNSEMEKVDGLLTCCSVFINK
 KIDS"
 360..659
 /misc_feature
 /notes="Amidnotransf; Region: Amidnotransferase. This
 family contains glycine (EC:2.1.4.1) and inosamine
 (EC:2.1.4.2) amidnotransferases, enzymes involved in
 creatine and streptomycin biosynthesis respectively. This
 family also includes arginine deaminases, EC:3.5.3.6.
 These enzymes catalyse the reaction: arginine + H2O ==>
 citrulline + NH3. Also found in this family is the
 Streptococcus anti tumour glycoprotein"
 /db_xref="CDD:pfam02274"
 ORIGIN
 Alignment Scores:
 Pred. NO.: 3.31e-59 Length: 3420
 Score: 756.00 Matches: 145
 Percent Similarity: 97.40% Conservative: 5
 Best Local Similarity: 94.16% Mismatches: 4
 Query Match: 51.99% Indels: 0
 DB: 10 Gaps: 0
 US-09-889-733B-2 (1-285) x BC060266 (1-3420)
 QY 132 LeuPheThrGlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnA-gglyVala 151
 Db 216 TTGCTCAGCGCAGAGATTTTGTGGGCTTCCAAAGAACAAATCAACGAGGTGCT 275
 QY 152 GluIleLeuAlaAspThrPheLysAspTyAlaValSerThrValProValAlaAspGly 171
 Db 276 GAAATCTGGCTGATCTCTTTAAGGACTACGAGTCTCTACAGTCCCTGGCGCGATCT 335
 QY 172 LeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySerSer 191
 Db 336 TTGCTTTTAAAGAGTTTCTGCAGCATGGCCGGAGCCCACTGATTGCAATAGGTCACG 395

QY 192 GluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyrAspLys 211
 Db |||||
 396 GAATCTGCACAGAAGCCCTCAAGATCATGCAACACAGATGAGTGACCATGTTATGACAAG 455
 QY 212 LeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGly 231
 Db |||||
 456 CTCACGTATACCCGACGACATGGCCGCAACTGCATATATCTAAATATCCCCAGCAAGGG 515
 QY 232 HisValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLysValTyrGluLys 251
 Db |||||
 516 CATGCTCTGCTGCACCGAACCCCGAGAGAGTACCCAGAAAGCGCAAGGTCTATGAGAA 575
 QY 252 LeuLysAspHisMetLeuIleProValSerMetSerGluLeuGluLysValAspGlyLeu 271
 Db |||||
 576 CTCAGGACCATCTACTGATCCCTGTGAGCAACTCGGAGATGAAAAGGTGACGGCTTG 635
 QY 272 LeuThrCysCysSerValLeuIleAsnLysLysValAspSer 285
 Db |||||
 636 CTCACCTGTGCTCCGTTTTTTATTAACRAGAGATAGACTCC 677

Search completed: June 8, 2004, 00:22:52
 Job time : 3219 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 20:32:55 ; Search time 3410 Seconds
(without alignments)
10905.667 Million cell updates/sec

Title: US-09-889-733B-1

Perfect score: 858

Sequence: 1 atggcggcctggccacc.....acaagaagtagactctcta 858

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.man.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	858	100.0	858	6	BD251986 Screen me
2	858	100.0	858	6	AX032835 Sequence
3	772	90.0	1624	9	BC033680 Homo sapi
4	772	90.0	1633	6	BD094073 Shear str
5	772	90.0	1633	9	AB001915 Homo sapi
6	772	90.0	4208	6	AX548046 Sequence
7	728	84.8	3737	9	BC043235 Homo sapi
8	544	63.4	3908	9	HSM808292
9	544	63.4	3908	9	AK098405 Homo sapi
10	232	27.2	2895	9	AK098405 Homo sapi
11	232	27.0	104071	2	AL162260 Homo sapi
12	232	27.0	158405	2	AC019261
13	232	27.0	168953	9	AL360219 Human DNA
14	148	17.2	83946	9	HS123M24
15	85	9.9	420	6	AX898477 Sequence
16	85	9.9	420	6	BD034010 Sequence
17	79	9.2	110000	2	Continuation (3 of
18	79	9.2	304208	2	AC14263_2
19	59	6.9	1120	10	BC034505
20	54	6.3	3008	6	AX683109
21	54	6.3	3008	10	D86041 Rattus norv
22	54	6.3	228940	2	AC126890
23	47	5.5	2289	10	AK128909 Mus muscu
24	47	5.5	3420	10	BC060286
25	44	5.1	212111	2	AC094482 Rattus no
26	43	5.0	626	11	BV064692 S212P6014
27	43	5.0	186434	2	AC126412 Mus muscu
28	43	5.0	200007	2	AC123684 Mus muscu
29	26	3.0	1344	5	BC056074 Xenopus l
30	25	2.9	58539	2	AC137903 Mus muscu
31	23	2.7	2520	10	AF127389 Rattus no
32	23	2.7	2526	10	AF301161 Mus muscu
33	23	2.7	2526	10	AF301162 Mus muscu
34	23	2.7	2771	6	BD233637 Nucleic a
35	23	2.7	237743	2	AC129831 Rattus no
36	22	2.6	1150	5	AB095027 Gallus ga
37	22	2.6	90348	1	AF497482 Micromono
38	21	2.4	98857	9	AC074097 Homo sapi
39	21	2.4	98935	9	AC004935 Homo sapi
40	21	2.4	162173	2	AC123569 Rattus no
41	21	2.4	181835	9	AC146227 Pan trogl
42	21	2.4	210408	2	AC142421 Rattus no
43	21	2.4	212026	2	AC146053 Pan trogl
44	21	2.4	247412	2	AC115249 Rattus no
45	20	2.3	921	11	BV005803 BARC0051

ALIGNMENTS

RESULT 1
BD251986
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BD251986
Screen method.
BD251986
BD251986.1 GI:33061756
JP 2002535001-A/1.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 858)
Vallance,P.J.T., Leiper,J.M., Whitley,G.S.J. and Charles,I.G.
Screen method
Patent: JP 2002535001-A 1 22-OCT-2002;

858 bp DNA linear PAT 17-JUL-2003

Pred. No. is the number of results predicted by chance to have a

UNIVERSITY COLLEGE LONDON
OS Homo sapiens (human)
PN JP 2002535001-A/1
PD 22-OCT-2002
PF 26-JAN-2000 JP 2000596131
PR 26-JAN-1999 GB 9901705.S.04-JUN-1999 GB 9913066.8 PI
PATRICK JOHN THOMPSON VALLANCE, JAMES MITCHELL LEIPER, GUY ST PI
JOHN WHITLEY,
PI IAN GEORGE CHARLES
PC C12N15/09, A01K67/027, A61K31/198, A61K38/00, A61K45/00, A61P3/06,
PC A61P9/02, A61P9/10, A61P9/12, A61P13/12, A61P25/06, A61P25/18, PC
PC A61P9/04, A61P9/10, A61P9/12, A61P13/12, A61P25/06, A61P25/18, PC
A61P25/28,
PC A61P29/00, A61P31/04, A61P35/00, C07K16/40, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12N9/78, C12Q1/02, C12Q1/34, G01N33/15, G01N33/50 PC
C12N15/00, C12N5/00,
PC A61K37/02
CC Screen method
FH Key
FT CDS Location/Qualifiers
1. .858
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

Query Match 100.0%; Score 858; DB 6; Length 858;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 858; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGCTCGGACACCCCTTCGGCTTCGGCGGCGCCACCCACCGCTGTGTGGGGG 60
DB 1 ATGGCGGCTCGGACACCCCTTCGGCTTCGGCGGCGCCACCCACCGCTGTGTGGGGG 60
QY 61 CTACCCGAGTCGCTTCGCCAGCAGCGCTGAGAGCGCCAGAGCGGAGGCTGACGCTC 120
DB 61 CTACCCGAGTCGCTTCGCCAGCAGCGCTGAGAGCGCCAGAGCGGAGGCTGACGCTC 120
QY 121 GCCCGCGGAGCGGACGACGCTCTACGTTGGGCGTGTGGGACGACGCTGTGGGGCTG 180
DB 121 GCCCGCGGAGCGGACGACGCTCTACGTTGGGCGTGTGGGACGACGCTGTGGGGCTG 180
QY 181 CAGGTGTGAGTCGCTTCGCCAGCAGCGCTTCGGCGGCGCCACCCACCGCTGTGTGGGGG 240
DB 181 CAGGTGTGAGTCGCTTCGCCAGCAGCGCTTCGGCGGCGCCACCCACCGCTGTGTGGGGG 240
QY 241 GCCGTGTGTGCGAGGAGCGGCGCTTCATCCCGACCGCGGCGCGCGGAGCGGAGG 300
DB 241 GCCGTGTGTGCGAGGAGCGGCGCTTCATCCCGACCGCGGCGCGCGGAGCGGAGG 300
QY 301 GAGTTGACATGATGAAGAGCATTAGAAAGCTTCACTCAATATAGTAGATGAAA 360
DB 301 GAGTTGACATGATGAAGAGCATTAGAAAGCTTCACTCAATATAGTAGATGAAA 360
QY 361 GATGAAATGCACTTTAGATGCGGAGATGTTTTTATTCACAGGACGAGATTTTTTGTG 420
DB 361 GATGAAATGCACTTTAGATGCGGAGATGTTTTTATTCACAGGACGAGATTTTTTGTG 420
QY 421 GGCTTTTCCAAAAGGACAAATCAACAGAGTCGTGAATCTTGGCTGATCTTTAAGGAC 480
DB 421 GGCTTTTCCAAAAGGACAAATCAACAGAGTCGTGAATCTTGGCTGATCTTTAAGGAC 480
QY 481 TATGCACTTCCACAGTCGCGAGTGGGCTGCTGATTTGAAGATTTCTGCAGCATG 540
DB 481 TATGCACTTCCACAGTCGCGAGTGGGCTGCTGATTTGAAGATTTCTGCAGCATG 540
QY 541 GCTGGGCTTAACCTGATCGAAATGAGTGTGATGAAATCTGCAAGAGGCGCTTTAAGATC 600
DB 541 GCTGGGCTTAACCTGATCGAAATGAGTGTGATGAAATCTGCAAGAGGCGCTTTAAGATC 600
QY 601 ATGCAACAGATGAGTGACCCCGCTACGACAACTCACGTGCTGCTGATCAGTAGCAGCA 660

601 ATGCAACAGATGAGTGACCCCGCTACGACAACTCACTGTGCTGATCAGTAGCAGCA 660
QY 661 AACTGTATATCTAATATATCCCAACCAAGGCGACGCTTGTGTGACCAACCCCGGAA 720
DB 661 AACTGTATATCTAATATATCCCAACCAAGGCGACGCTTGTGTGACCAACCCCGGAA 720
QY 721 GAGTATCCAGAAAGTGCAGAGTTTATGAGAACTGAAGACCATATGCTGATCCCGTG 780
DB 721 GAGTATCCAGAAAGTGCAGAGTTTATGAGAACTGAAGACCATATGCTGATCCCGTG 780
QY 781 AGCATGCTGAATGGAAGAGTGGATGGCTGCTCCTACCTGCTGCTCAGTTTAAATTAAC 840
DB 781 AGCATGCTGAATGGAAGAGTGGATGGCTGCTCCTACCTGCTGCTCAGTTTAAATTAAC 840
QY 841 AAGAAGGTAGACTCCTGA 858
DB 841 AAGAAGGTAGACTCCTGA 858

RESULT 2
AX032835
LOCUS
DEFINITION
Sequence 1 from Patent WO0044888.
AX032835
ACCESSION
AX032835.1 GI:10279809
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Leiper, J.M., Whitley, G.S., Charles, I.G. and Vallance, P.J.
Screen method
Patent: WO 0044888-A 1 03-AUG-2000;
LEIPER JAMES MITCHELL (GB); WHITLEY GUY ST JOHN (GB); UNIV LONDON
(GB); CHARLES IAN GEORGE (GB); VALLANCE PATRICK JOHN THOMPSON
(GB)

FEATURES
Location/Qualifiers
1. .858
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1. .858
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC09947.1"
/db_xref="GI:10279810"
/db_xref="REMBL:CAC09947"
/translation="MAGLHPSAFGRATHAVVRLPESLQHALRSKAEVDVARAE
RQHLVYVGLSKGLQVELPADSLPDCVFDVAVVCEETALIRPAPRRKEV
DMKEALEKLQNLIVEMKDNATLDGDFVLTGREFFVGLSKRTNQSGILATPDDI
YAVSTVPVADGLHLKSPCSMAGPNLIIAGSSAQKALIMQMSDRYDKLTPDDI
AANCYLNIPNKGHVLHRTPEEYPSAKVYKLDHMLIPVMSLEKVDGLLTCCS
VLINKYDS"

ORIGIN
Query Match 100.0%; Score 858; DB 6; Length 858;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 858; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGGCTCGGACACCCCTTCGGCTTCGGCGGCGCCACCCACCGCTGTGTGGGGG 60
DB 1 ATGGCGGCTCGGACACCCCTTCGGCTTCGGCGGCGCCACCCACCGCTGTGTGGGGG 60
QY 61 CTACCCGAGTCGCTTCGCCAGCAGCGCTGAGAGCGCCAGAGCGGAGGCTGACGCTC 120
DB 61 CTACCCGAGTCGCTTCGCCAGCAGCGCTGAGAGCGCCAGAGCGGAGGCTGACGCTC 120
QY 121 GCCCGCGGAGCGGACGACGCTCTACGTTGGGCGTGTGGGACGACGCTGTGGGGCTG 180
DB 121 GCCCGCGGAGCGGACGACGCTCTACGTTGGGCGTGTGGGACGACGCTGTGGGGCTG 180

family contains glycine (EC:2.1.4.1) and inosamine (EC:2.1.4.2) amidotransferases, enzymes involved in creatine and streptomycin biosynthesis respectively. This family also includes arginine deiminases, EC:3.5.3.6. These enzymes catalyze the reaction: arginine + H₂O <=> citrulline + NH₃. Also found in this family is the Streptococcus anti tumor glycoprotein"
/db_xref="CDD:pfam02274"

```
ORIGIN
Query Match      90.0%; Score 772; DB 9; Length 1624;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 822; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 23 CGCCTTCGGCGGCGCCACCCAGCCGCTGGTGGCGCGCTACCCGAGTCGCTGCCAGC 82
Db 122 CGCGCTTCGGCGGCGCCACCCAGCCGCTGGTGGCGCGCTACCCGAGTCGCTGCCAGC 181
QY 83 AGCGCTCAGAGAGCCCAAGGCGGAGGAGTGGAGTCCGCCGCGCGCAACCGCAGACCC 142
Db 182 AGCGCTCAGAGAGCCCAAGGCGGAGGAGTGGAGTCCGCCGCGCGCAACCGCAGACCC 241
QY 143 AGCTCTACGTGGCGCTGTGGCGCAAGCTGGGCTCAGGTGGTGGAGTGGCGGCGG 202
Db 242 AGCTCTACGTGGCGCTGTGGCGCAAGCTGGGCTCAGGTGGTGGAGTGGCGGCGG 301
QY 203 ACGGAGCGCTTCGGAGCTGCGTTCGCTGGAGGAGTGGCGGCTGGTGGCGAGGAGCG 262
Db 302 ACGGAGCGCTTCGGAGCTGCGTTCGCTGGAGGAGTGGCGGCTGGTGGCGAGGAGCG 361
QY 263 CCCTCATCACCCGACCGCGGCGCGAGCGGAGGAGGAGTGGATGATGAAAGAG 322
Db 362 CCCTCATCACCCGACCGCGGCGCGAGCGGAGGAGGAGTGGATGATGAAAGAG 421
QY 323 CATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAGATGMAATGCACTTTAGATG 382
Db 422 CATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAGATGMAATGCACTTTAGATG 481
QY 383 GCGGAGATGTTTATTCACAGCAGAGAAATTTTGTGGGCTTTTCCAAAAGGACAAATC 442
Db 482 GCGGAGATGTTTATTCACAGCAGAGAAATTTTGTGGGCTTTTCCAAAAGGACAAATC 541
QY 443 AACGAGGTGCTGAATCTTGGCTGATCTTTTAAAGGATGTCAGTCTCCAGTGGCAG 502
Db 542 AACGAGGTGCTGAATCTTGGCTGATCTTTTAAAGGATGTCAGTCTCCAGTGGCAG 601
QY 503 TGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCTAACCTGATCGCAA 562
Db 602 TGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCTAACCTGATCGCAA 661
QY 563 TTGGGTCTAGTGAATCTGCACAGAGGCGCTTTAAGATCATGCACAGATGATGACCAAC 622
Db 662 TTGGGTCTAGTGAATCTGCACAGAGGCGCTTTAAGATCATGCACAGATGATGACCAAC 721
QY 623 GCTACGACAACTCACTGTCCTGATGATACATAGCAGCAAACTGTATATCTTAATATCC 682
Db 722 GCTACGACAACTCACTGTCCTGATGATGATGATGATGATGATGATGATGATGATGAT 781
QY 683 CCAACAAAGGCGCGCTTTGCTGCAACCGAACCCCGAAGAGATATCCAGAAAGTGCAGAG 742
Db 782 CCAACAAAGGCGCGCTTTGCTGCAACCGAACCCCGAAGAGATATCCAGAAAGTGCAGAG 841
QY 743 TTTATCAGAACTGAGGACCATATGCTGATCCCGGTGGGATGCTGCACTGGAAGAG 802
Db 842 TTTATCAGAACTGAGGACCATATGCTGATCCCGGTGGGATGCTGCACTGGAAGAG 901
QY 803 TGGATGGGCTGCTCACTGCTGCTCAGTTTATTTAATTAACAAGAA 845
Db 902 TGGATGGGCTGCTCACTGCTGCTCAGTTTATTTAATTAACAAGAA 944
```

RESULT 4
BD094073

```
LOCUS      BD094073      1633 bp      DNA      linear      PAT 27-AUG-2002
DEFINITION Shear stress-responsive DNAs.
ACCESSION   BD094073
VERSION     BD094073.1  GI:22639661
KEYWORDS    WO 0125427-A/34.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1633)
AUTHORS     Nojima,H., Yoshisue,H., Obayashi,M., Ota,T., Kawabata,A.,
            Sakurada,K., Kuga,T., Sekine,S., Nakamura,Y. and Sugan,S.
TITLE       Shear stress-responsive DNAs
JOURNAL     Patent: WO 0125427-A 34 12-APR-2001;
            KIOWA HAKKO KOSYO CO LTD,HIROSHI NOJIMA,HAJIME YOSHISUE, MASAYA
            OBAAYASHI, TOSHIO OTA,AYAKO KAWABATA,KAZUHIRO SAKURADA,TEISURO KUGA,
            SUSUMU SEKINE, YUSUKE NAKAMURA, SUMIO SUGANO
COMMENT     OS Homo sapiens (human)
            PN WO 0125427-A/34
            PD 12-APR-2001
            PE 02-OCT-2000 WO 2000JP006840
            PF 01-OCT-1999 JP 99P 280976
            PR 01-OCT-1999 JP 99P 280976
            PI HIROSHI NOJIMA,HAJIME YOSHISUE,MASAYA OBAAYASHI,TOSHIO OTA, PI
            AYAKO KAWABATA,
            PI KAZUHIRO SAKURADA,TEISURO KUGA,SUSUMU SEKINE,YUSUKE NAKAMURA,
            SUMIO SUGANO
            PC C12N15/12,C07K14/435,C07K16/18,C12P21/02,C12Q1/68,A61K38/00,
            A61K48/00,A61P9/10,G01N33/50,G01N33/53
            PC A61K39/395,
            CC A61K48/00,A61P9/10,G01N33/50,G01N33/53
FEATURES     Location/Qualifiers
             1..1633
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
ORIGIN
Query Match      90.0%; Score 772; DB 6; Length 1633;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 822; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 23 CGCCTTCGGCGGCGCCACCCAGCCGCTGGTGGCGGCGCTACCCGAGTCGCTGCCAGC 82
Db 345 CGCCTTCGGCGGCGCCACCCAGCCGCTGGTGGCGGCGCTACCCGAGTCGCTGCCAGC 404
QY 83 ACGCGCTGAGAGCGCAAGGGCGAGGAGTGGACGTCGCCCGCGCGGAACCGCAGCACC 142
Db 405 ACGCGCTGAGAGCGCAAGGGCGAGGAGTGGACGTCGCCCGCGCGGAACCGCAGCACC 464
QY 143 AGCTCTACGTGGCGCTGTGGCGCAAGCTGGGCGTGCAGGTGGTGGAGTGGCGGCGG 202
Db 465 AGCTCTACGTGGCGCTGTGGCGCAAGCTGGGCGTGCAGGTGGTGGAGTGGCGGCGG 524
QY 203 ACGAGAGCGCTTCGGGACTGCTCTTCGTGGAGAGCGTGGCGGCTGGTGGAGGAGCGG 262
Db 525 ACGAGAGCGCTTCGGGACTGCTCTTCGTGGAGAGCGTGGCGGCTGGTGGAGGAGCGG 584
QY 263 CCTCTCATCCCGACCCCGGGCGCCGAGCCGCGAGGAAGAGGTTGACATGATGAAGAAG 322
Db 585 CCTCTCATCCCGACCCCGGGCGCCGAGCCGCGAGGAAGAGGTTGACATGATGAAGAAG 644
QY 323 CATTAGAAAACTTCAGCTCAATATAGTAGATGAAGATGAAGATGAAGATGAAGATGAAG 382
Db 645 CATTAGAAAACTTCAGCTCAATATAGTAGATGAAGATGAAGATGAAGATGAAGATGAAG 704
QY 383 GCGGAGATGTTTATTCACAGGCGAGAAATTTTGTGGGCTTTTCCAAAAGGACAAATC 442
Db 705 GCGGAGATGTTTATTCACAGGCGAGAAATTTTGTGGGCTTTTCCAAAAGGACAAATC 764
QY 443 ACGAGGTGCTGAATCTTGGCTGATCTTTTAAAGGACTATGCACTCCAGTCCAGTCCAG 502
```

Db 765 AACGAGTGCTGAAATCTTGGCTGATATCTTTAAGGACTATGAGTCTCCACAGTCCAG 824

QY 503 TGGCAGATGGGTTGCAATTCGAAGAGTTTCTGCAGCATGGCTGGGCTCAACCTGATCGCAA 562

Db 825 TGGCAGATGGGTTGCAATTCGAAGAGTTTCTGCAGCATGGCTGGGCTCAACCTGATCGCAA 884

QY 563 TTGGGCTAGTGAATCTGCACAGAGCCCTTAAAGATCATGCAACAGATGAGTGACCAAC 622

Db 885 TTGGGCTAGTGAATCTGCACAGAGCCCTTAAAGATCATGCAACAGATGAGTGACCAAC 944

QY 623 GCTACGCAAAACTCACTGTGCTGATGACATAGCAGCAAACTGTATATATCTAAATATCC 682

Db 945 GCTACGCAAAACTCACTGTGCTGATGACATAGCAGCAAACTGTATATATCTAAATATCC 1004

QY 683 CCAACAAAGGCGAGCTTCTGTCGACCGAACCCCGGAGAGTATCCAGAAAGTGCAAAAG 742

Db 1005 CCAACAAAGGCGAGCTTCTGTCGACCGAACCCCGGAGAGTATCCAGAAAGTGCAAAAG 1064

QY 743 TTTATGAGAAACTGAAGACCATATGCTGATCCCGTGAGCATGTCTGAACTGCAAAAG 802

Db 1065 TTTATGAGAAACTGAAGACCATATGCTGATCCCGTGAGCATGTCTGAACTGCAAAAG 1124

QY 803 TGGATGGGCTGCTCACCTGTGCTGCTCAGTTTAAATTAACAAGAA 845

Db 1125 TGGATGGGCTGCTCACCTGTGCTGCTCAGTTTAAATTAACAAGAA 1167

RESULT 5

AB001915

LOCUS

DEFINITION Homo sapiens mRNA for NG,NG-dimethylarginine dimethylaminohydrolase, complete cds.

ACCESSION AB001915

VERSION AB001915.1 GI:4160665

KEYWORDS NG,NG-dimethylarginine dimethylaminohydrolase.

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (sites)

AUTHORS Kimoto,M., Miyatake,S., Sasagawa,T., Yamashita,H., Okita,M., Oka,I., Ogawa,I. and Tsuji,H.

TITLE Purification, cDNA cloning and expression of human NG,NG-dimethylarginine dimethylaminohydrolase

JOURNAL Eur. J. Biochem. 258 (2), 863-868 (1998)

MEDLINE 98089653

PUBMED 9874257

REFERENCE 2 (bases 1 to 1633)

AUTHORS Kimoto,M.

TITLE Direct Submission

JOURNAL Submitted (15-MAR-1997) Masumi Kimoto, Okayama Prefectural University, Faculty of Health and Welfare Science, Department of Nutritional Science, Kuboki 111, Soja, Okayama 719-1197, Japan (E-mail:kimoto@fhw.oka-pu.ac.jp, Tel:0866-94-2158, Fax:0866-94-2158)

FEATURES

source

1. 1633

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="Kidney"

323..1180

/EC_number="3.5.3.18"

/codon_start=1

/product="NG,NG-dimethylarginine dimethylaminohydrolase"

/protein_id="BAA37117.1"

/db_xref="GI:4160666"

/translation="MAGLGPAPAFGATHAVVRLPESLGOHALPSAKGEVVDVARAE RQHLVGVLSKLGQVVELPADESLPCVFVEDVAVVCEBTALITRPGASRKEV DMKALEKQLNIVEMKENATLDGDFVFGLEFFGLSKTRQGRHAILADTFKD YAVSTVPVADGLHKLFCNSAGNENLAISSSAQKLMQMSDHRDKLTVPDDI AANCYININPKGHVLLHRTPEEPYPSAKVYKLEKLDHMLIPVMSLEKVDLLCCS VLINKVDS"

CDS

ORIGIN

Query Match 90.0%; Score 772; DB 9; Length 1633;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 822; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 23 CGCCCTTCGGCGGGGCGACCCACCGCGCTGTGGTGTGGGGCGCTACCCGAGTCGCTCTGCCAGC 82

Db 345 CGCCCTTCGGCGGGGCGACCCACCGCGCTGTGGTGTGGGGCGCTACCCGAGTCGCTCTGCCAGC 404

QY 83 ACGCCCTCAGAAAGCCCAAGGGCGAGAGGTGGAGCTGCCCGCCGCGGAAAGGGCAGACACC 142

Db 405 ACGCCCTCAGAAAGCCCAAGGGCGAGAGGTGGAGCTGCCCGCCGCGGAAAGGGCAGACACC 464

QY 143 AGCTCTACGTGGGCGTGTCTGGGCGAGCAAGCTGGGGCTGCAGGTGGTGGAGCTGCCCGCGG 202

Db 465 AGCTCTACGTGGGCGTGTCTGGGCGAGCAAGCTGGGGCTGCAGGTGGTGGAGCTGCCCGCGG 524

QY 203 ACGAGAGCGCTTCGGGAGCTGCGTCTTCGTGGAGGAGCGTGGCCGCTGTGGTGGAGAGACGG 262

Db 525 ACGAGAGCGCTTCGGGAGCTGCGTCTTCGTGGAGGAGCGTGGCCGCTGTGGTGGAGAGACGG 584

QY 263 CCTCATCACCCGAGCCCGGGCGCGAGCGCGGAGGAGGTTGACATGATGAAGAAG 322

Db 585 CCTCATCACCCGAGCCCGGGCGCGAGCGCGGAGGAGGTTGACATGATGAAGAAG 644

QY 323 CATTAGAABAACTTCAGCTCAATATAGTAGAGTGAAGTGAABAACTTCAGCTTTAGATG 382

Db 645 CATTAGAABAACTTCAGCTCAATATAGTAGAGTGAAGTGAABAACTTCAGCTTTAGATG 704

QY 383 GCGGAGATGTTTATTCACAGGCGAGAGAAATTTTGTGGGCTTTCCAAAAGGCAAAATC 442

Db 705 GCGGAGATGTTTATTCACAGGCGAGAGAAATTTTGTGGGCTTTCCAAAAGGCAAAATC 764

QY 443 AACGAGGTGCTGAATCTTTGGCTGATCTTTAAGGACTATGAGTATGAGTATGAGTATGAGT 502

Db 765 AACGAGGTGCTGAATCTTTGGCTGATCTTTAAGGACTATGAGTATGAGTATGAGTATGAGT 824

QY 503 TGGCAGATGGGTTGCAATTCGAAGAGTTTCTGCAGCATGGCTGGGCTAACTGATCGCAA 562

Db 825 TGGCAGATGGGTTGCAATTCGAAGAGTTTCTGCAGCATGGCTGGGCTAACTGATCGCAA 884

QY 563 TTGGGCTAGTGAATCTGCACAGAGCCCTTAAAGATCATGCAACAGATGAGTGACCAACC 622

Db 885 TTGGGCTAGTGAATCTGCACAGAGCCCTTAAAGATCATGCAACAGATGAGTGACCAACC 944

QY 623 GCTACGCAAAACTCACTGTGCTGATGACATAGCAGCAAACTGTATATATCTAAATATCC 682

Db 945 GCTACGCAAAACTCACTGTGCTGATGACATAGCAGCAAACTGTATATATCTAAATATCC 1004

QY 683 CCAACAAAGGCGAGCTTCTGTCGACCGAACCCCGGAGAGTATCCAGAAAGTGCAAAAG 742

Db 1005 CCAACAAAGGCGAGCTTCTGTCGACCGAACCCCGGAGAGTATCCAGAAAGTGCAAAAG 1064

QY 743 TTTATGAGAAACTGAAGACCATATGCTGATCCCGTGAGCATGTCTGAACTGCAAAAG 802

Db 1065 TTTATGAGAAACTGAAGACCATATGCTGATCCCGTGAGCATGTCTGAACTGCAAAAG 1124

QY 803 TGGATGGGCTGCTCACCTGTGCTGCTCAGTTTAAATTAACAAGAA 845

Db 1125 TGGATGGGCTGCTCACCTGTGCTGCTCAGTTTAAATTAACAAGAA 1167

RESULT 6

AX548046

LOCUS

DEFINITION Sequence 22 from Patent WO02066654.

ACCESSION AX548046

VERSION AX548046.1 GI:25813142

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Astromoff, A., Au-Young, J., Baughn, M.R., Ding, L., Duggan, B.M.,
Forsythe, I.J., Gietzen, K.J., Griffin, J.A., Lee, E.A., Lu, Y.,
Richardson, T.W., Ring, H.Z., Sanjanwala, M.X., Swannakar, A.,
Wallia, N.K., Warren, S.A., Xu, Y., Yue, H. and Zebajjadian, Y.
Drug metabolizing enzymes
Patent: WO 02066654-A 22 29-AUG-2002;
Incyte Genomics, Inc. (US)
Location/Qualifiers
1. 4208
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7671089CBI"

ORIGIN
Query Match 90.0%; Score 772; DB 6; Length 4208;
Best Local Similarity 99.9%; Pred No 0;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 23 CGCGCTTCGGCGCCGACCCACGCGCGTGGTGGCGCGCTCCGAGTGGTGGCGCGC 82
Db 390 CGCGCTTCGGCGCCGACCCACGCGCGTGGTGGCGCGCTCCGAGTGGTGGCGCGC 449

QY 83 AGCGCTGAGAGCCGCGAGGCGGAGGAGTGGAGTGGCGCGCGCGCGCGCGCGCGC 142
Db 450 AGCGCTGAGAGCCGCGAGGCGGAGGAGTGGAGTGGCGCGCGCGCGCGCGCGCGC 509

QY 143 AGCTCTACGTGGCGCGTGGTGGCGAGCAAGCTGGGCGCTCAGGTGGTGGCGCGCGC 202
Db 510 AGCTCTACGTGGCGCGTGGTGGCGAGCAAGCTGGGCGCTCAGGTGGTGGCGCGCGC 569

QY 203 ACAGAGCGCTTCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 262
Db 570 ACAGAGCGCTTCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 629

QY 263 CCCTCATACCGC 322
Db 630 CCCTCATACCGC 689

QY 323 CATTTAGAAAACCTCAGCTCAATATAGTAGAGATGAAGATGAAGATGAAGATGAAG 382
Db 690 CATTTAGAAAACCTCAGCTCAATATAGTAGAGATGAAGATGAAGATGAAGATGAAG 749

QY 383 GCGGAGATGTTTATTTACAGGACAGAAATTTTGTGGCGCTTTTCCAAAAGGACAAATC 442
Db 750 GCGGAGATGTTTATTTACAGGACAGAAATTTTGTGGCGCTTTTCCAAAAGGACAAATC 809

QY 443 AACGAGGTGCTGAATCTTGGCTGATATTTTAAAGCAATATGCAATATGCAATATGCA 502
Db 810 AACGAGGTGCTGAATCTTGGCTGATATTTTAAAGCAATATGCAATATGCAATATGCA 869

QY 503 TGGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 562
Db 870 TGGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 929

QY 563 TTGGGTCTAGTGAATCTGACAGAGGCGCTTTAAGATCATGCAACAGATGATGACCAAC 622
Db 930 TTGGGTCTAGTGAATCTGACAGAGGCGCTTTAAGATCATGCAACAGATGATGACCAAC 989

QY 623 GCTAGCAGAACTCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 682
Db 990 GCTAGCAGAACTCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1049

QY 683 CCAACAAAGGCGCGCTTGTGTGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 742
Db 1050 CCAACAAAGGCGCGCTTGTGTGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1109

QY 743 TTTATGAGAACTGAAGACCAATATGCTGATCCCGTGGAGATGCTGAACTGAAAGAGG 802
Db 1110 TTTATGAGAACTGAAGACCAATATGCTGATCCCGTGGAGATGCTGAACTGAAAGAGG 1169

QY 903 TGGATGGGTGCTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 845
Db 1170 TGGATGGGTGCTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1212

RESULT 7
BC043235 3737 bp mRNA linear PRI 30-SEP-2003
LOCUS Homo sapiens dimethylarginine dimethylaminohydrolase 1, mRNA (cdna
DEFINITION clone IMAGE:5295723), partial cds.
ACCESSION BC043235.2 GI:34783628
VERSION BC043235
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3737)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raja, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S.J., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E.,
Schnerk, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3737)
Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
On Sep 16, 2003 this sequence version replaced gi:28175756.
Contact: MGC help desk
Email: cgabs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeay, Steven
Ness, Pawan Pandon, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Small, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 76 Row: j Column: 15.
Location/Qualifiers
1. 3737

FEATURES
source

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5295723"
/tissue_type="Testis"
/clone_lib="NIH MGC 97"
/lab_host="DH10B"
/notes="Vector: pBluescript"
<1..3737
/gene="DDAH1"
/notes="synonym: DDAH"
/db_xref="LocusID:23576"
/db_xref="MIM:604743"
<1..741
/codon_start=1
/product="DDAH1 protein"
/protein_id="AA43235.2"
/db_xref="GI:34783629"
/db_xref="LocusID:23576"
/db_xref="MIM:604743"
/translations="VARASRQHLYVGLSKLQVLPADSLPDCVFVEDVAVV
CELTALTRPGAPSRKEVDMKEALEKLQNLIVEMKDENATLDGQDLVLTGRERFVG
LSKRTNORGAILADTFKQYAVSTVPVADGLHLKSPCSNAGPNLIAIGSSESAQALK
IMQMSDHRVYDKLTVPDIDIAAHCIIYINPKGHVLLHRTPEYPESAKVYELKDHML
IPVMSLEKVDGDLITCCSVLLINKKVD"
misc_feature
1..720
/gene="DDAH1"
/notes="Amidinoetransf; Region: Amidinotransferase. This
family contains glycine (EC:2.1.4.1) and inosamine
(EC:2.1.4.2) amidinotransferases, enzymes involved in
creatine and streptomycin biosynthesis respectively. This
family also includes arginine deiminases, EC:3.5.3.6.
These enzymes catalyze the reaction: arginine + H2O <=>
citrulline + NH3. Also found in this family is the
Streptococcus anti tumor glycoprotein"
/db_xref="CDD:pfam02274"

ORIGIN
Query Match 84.8%; Score 728; DB 9; Length 3737;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 GTGCGCCGCGGAAACGGCAGCACCAGCTCTACGTGGCGGTGCTGGCGAGCAGCTGGGG 177
DB 1 GTGCGCCGCGGAAACGGCAGCACCAGCTCTACGTGGCGGTGCTGGCGAGCAGCTGGGG 60

QY 178 CTGCAAGTGTGGAGCTGCCGCGCAGCAGAGCCTTCCGAGCTGCCTTCGTGGAGGAC 237
DB 61 CTGCAAGTGTGGAGCTGCCGCGCAGCAGAGCCTTCCGAGCTGCCTTCGTGGAGGAC 120

QY 238 GTGCGCTGTGTGCGAGGAGACGGCCCTCATCACCAGCCCGGGCGCGAGCGGAGG 297
DB 121 GTGCGCTGTGTGCGAGGAGACGGCCCTCATCACCAGCCCGGGCGCGAGCGGAGG 180

QY 298 AAGGAGGTTCACATGATGAAGAAGACATTAGAAAACCTTCAGCTCAATATAGTAGAGATG 357
DB 181 AAGGAGGTTCACATGATGAAGAAGACATTAGAAAACCTTCAGCTCAATATAGTAGAGATG 240

QY 358 AAGATGAAATCAACTTTAGATGCGGAGATGTTTATTCACAGCAGAGATTTTTT 417
DB 241 AAGATGAAATCAACTTTAGATGCGGAGATGTTTATTCACAGCAGAGATTTTTT 300

QY 418 GTGCGCCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAATCTTGGCTGATCTTTAAG 477
DB 301 GTGCGCCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAATCTTGGCTGATCTTTAAG 360

QY 478 GACTATGACGTCTCCACAGTGCAGTGGCAGATGGGTTCATTTGAAGAGTTTCTCAGC 537
DB 361 GACTATGACGTCTCCACAGTGCAGTGGCAGATGGGTTCATTTGAAGAGTTTCTCAGC 420

QY 538 ATGGCTGGGCTAACCTGATGCGAATTTGGTCTAGTGAATCTGCACAGAGGCCCTTAG 597
DB 362 ATGGCTGGGCTAACCTGATGCGGAGATGTTTATTCACAGGAGAGATTTTGTGG 421
```

```
DB 421 ATGGCTGGGCTAACCTGATGCGAATTTGGTCTAGTGAATCTGCACAGAGGCCCTTAAG 480
QY 598 ATCATGCAACAGATGAGTGACCGCTACGACCAAACTCACTGTGCTGATGACATAGCA 657
DB 481 ATCATGCAACAGATGAGTGACCGCTACGACCAAACTCACTGTGCTGATGACATAGCA 540
QY 658 GCAAACTGTATATCTATAATATCCCAACAAAGGCGACGCTTCTGCTCACCGAACCCCG 717
DB 541 GCAAACTGTATATCTATAATATCCCAACAAAGGCGACGCTTCTGCTCACCGAACCCCG 600
QY 718 GAAGAGTATCCAGAAAGTGCAAGGTTTATGAGAACTGGAAGGACCATATGTCATCCC 777
DB 601 GAAGAGTATCCAGAAAGTGCAAGGTTTATGAGAACTGGAAGGACCATATGTCATCCC 660
QY 778 GTGAGCATGTCTGAACCTGCAAAAGGTGATGGGTGCTGCTGCTGCTGCTGCTGCTGCT 837
DB 661 GTGAGCATGTCTGAACCTGCAAAAGGTGATGGGTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 838 AACAGAA 845
DB 721 AACAGAA 728

RESULT 8
HSM808292
LOCUS HSM808292 3908 bp mRNA linear PRI 30-AUG-2003
DEFINITION Homo sapiens mRNA; cDNA DKFZp686N2176 (from clone DKFZp686N2176).
ACCESSION BX648145
VERSION BX648145.1 GI:34367304
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3908)
AUTHORS Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
CONSTRM The German Human cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686N2176) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
source
1..3908
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686N2176"
/tissue_type="human endometrium carcinoma cell line"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiI + SfiII"
/dev_stage="adult"

ORIGIN
Query Match 63.4%; Score 544; DB 9; Length 3908;
Best Local Similarity 100.0%; Pred. No. 2.9e-283;
Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 AGTTGACATGATGAAGAAGCATTAGAAAACCTTCAGCTCAATATAGTAGAGTGAAG 361
DB 382 AGTTGACATGATGAAGAAGCATTAGAAAACCTTCAGCTCAATATAGTAGAGTGAAG 441
QY 362 ATGAAATGCAACTTTAGATGCGGAGATGTTTATTCACAGGAGAGATTTTGTGG 421
```


* This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 4008: contig of 4008 bp in length
 4108: gap of 100 bp
 4109 17620: contig of 13512 bp in length
 17621 17720: gap of 100 bp
 17721 29332: contig of 11612 bp in length
 29333 29432: gap of 100 bp
 29433 34595: contig of 5163 bp in length
 34596 34695: gap of 100 bp
 34696 51023: contig of 16328 bp in length
 51024 51123: gap of 100 bp
 51124 56471: contig of 5348 bp in length
 56472 56571: gap of 100 bp
 56572 62357: contig of 5786 bp in length
 62358 62457: gap of 100 bp
 62458 104071: contig of 41614 bp in length.

FEATURES

source

1..104071
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="p22.3-31.1"
 /clone="RP4-611J7"
 /clone_lib="RPC1-4"

misc_feature

1..4008
 /notes="assembly fragment:00515
 fragment chain:1
 clone end:SP6
 vector side:left"
 4109..17620
 /notes="assembly fragment:00727
 fragment chain:1"
 17721..29332
 /notes="assembly fragment:00262
 fragment chain:1"
 29433..34595
 /notes="assembly fragment:00257
 fragment chain:1"
 34696..51023
 /notes="assembly fragment:00672
 fragment chain:1"
 51124..56471
 /notes="assembly fragment:00205"
 56572..62357
 /notes="assembly fragment:01056"
 62458..104071
 /notes="assembly fragment:00375
 clone end:T7
 vector side:right"

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

ORIGIN

Query Match 27.0%; Score 232; DB 2; Length 104071;
 Best Local Similarity 99.6%; Pred. No. 5.3e-114;
 Matches 282; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 23 CCGCTTTCCGCGCGGACCCACCGCGTGTGTGGCGGCTTACCCGAGTGCCTCTGCCAGC 82
 DB 37550 CCGCTTTCCGCGCGGACCCACCGCGTGTGTGGCGGCTTACCCGAGTGCCTCTGCCAGC 37491
 QY 83 ACGCGTGTGAGAACGCGCAACGCGGAGAGTGTGACGTTCGCCCGCGGAAACGCGACACC 142
 DB 37490 ACGCGTGTGAGAACGCGCAACGCGGAGAGTGTGACGTTCGCCCGCGGAAACGCGACACC 37431
 QY 143 AGCTTACGTGGCGTGTCTGGGAGAGCAAGTGTGGGTGTGGAGTGTGGAGTGTGGCGCGC 202
 DB 37430 AGCTTACGTGGCGTGTCTGGGAGAGCAAGTGTGGGTGTGGAGTGTGGAGTGTGGCGCGC 37371
 QY 203 ACAGAGACCTTCCGAGTCCGCTCTTTCGTGGAGACGTGCGCGGTGTGTCCGAGGAGACGG 262
 DB 37370 ACAGAGACCTTCCGAGTCCGCTCTTTCGTGGAGACGTGCGCGGTGTGTCCGAGGAGACGG 37311

QY 263 CCCTCATCCCGACCCCGT 305
 DB 37310 CCCTCATCCCGACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 37268

RESULT 11

AC019261

LOCUS

DEFINITION

AC019261

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC019261 158405 bp DNA linear HTG 12-MAR-2000
 Homo sapiens clone RP11-2P7, WORKING DRAFT SEQUENCE, 38 unordered
 pieces.

AC019261.3 GI:7229998

HTG; HTGS_PHASE1; HTGS_DRAFT.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 158405)

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

Homo sapiens, clone RP11-2P7

Unpublished

2 (bases 1 to 158405)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Becker, R., Bada, F.,
 Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G., Castle, A.,
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
 Dearrellano, K., Dewar, K., Domino, M., Doyle, N., Fenestor, J.,
 Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,
 Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
 Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K.,
 MacDonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
 McPheeters, R., Meldrim, J., Meneus, L., Morrow, J., Naylor, J.,
 Norman, C. H., O'Connor, R., O'Donnell, P., Olivari, T. M., Peterson, K.,
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rottman, D.,
 Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Tixell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
 Zimmer, A. and Zody, M.

Direct Submission

Submitted (31-DEC-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 12, 2000 this sequence version replaced gi:6899774.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L2694

Center clone name: 2.P.7

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 135380 bases at least Q40

Consensus quality: 146343 bases at least Q30

Consensus quality: 150850 bases at least Q20

Insert size: 173000; agarose-fp

Insert size: 154705; sum-of-contigs

Quality coverage: 2.8 in Q20 bases; agarose-fp

Quality coverage: 3.2 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 38 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will
* be preserved.
* 1 1607: contig of 1607 bp in length
* 1608 1707: gap of 100 bp
* 1708 3028: contig of 1321 bp in length
* 3029 3128: gap of 100 bp
* 3129 4889: contig of 1761 bp in length
* 4890 4989: gap of 100 bp
* 4990 6433: contig of 1444 bp in length
* 6434 6533: gap of 100 bp
* 6534 8085: contig of 1552 bp in length
* 8086 8185: gap of 100 bp
* 8186 9968: contig of 1783 bp in length
* 9969 10068: gap of 100 bp
* 10069 11857: contig of 1789 bp in length
* 11858 11957: gap of 100 bp
* 11958 14371: contig of 2414 bp in length
* 14372 14471: gap of 100 bp
* 14472 16047: contig of 1576 bp in length
* 16048 16147: gap of 100 bp
* 16148 18676: contig of 2529 bp in length
* 18677 18776: gap of 100 bp
* 18777 21785: contig of 3009 bp in length
* 21786 21885: gap of 100 bp
* 21886 24973: contig of 3088 bp in length
* 24974 25073: gap of 100 bp
* 25074 27441: contig of 2368 bp in length
* 27442 30718: contig of 3177 bp in length
* 30719 30818: gap of 100 bp
* 30819 33922: contig of 3174 bp in length
* 33923 34092: gap of 100 bp
* 34093 36731: contig of 2639 bp in length
* 36732 36831: gap of 100 bp
* 36832 39283: contig of 2452 bp in length
* 39284 39383: gap of 100 bp
* 39384 41838: contig of 2455 bp in length
* 41839 41938: gap of 100 bp
* 41939 44209: contig of 2271 bp in length
* 44210 44309: gap of 100 bp
* 44310 47729: contig of 3420 bp in length
* 47730 47829: gap of 100 bp
* 47830 50393: contig of 2564 bp in length
* 50394 50493: gap of 100 bp
* 50494 56065: contig of 5572 bp in length
* 56066 56165: gap of 100 bp
* 56166 61307: contig of 5141 bp in length
* 61308 61406: gap of 100 bp
* 61407 65295: contig of 3889 bp in length
* 65296 65395: gap of 100 bp
* 65396 71547: contig of 6152 bp in length
* 71548 71647: gap of 100 bp
* 71648 77172: contig of 5525 bp in length
* 77173 77272: gap of 100 bp
* 77273 80143: contig of 2871 bp in length
* 80144 80243: gap of 100 bp
* 80244 84307: contig of 4064 bp in length
* 84308 84407: gap of 100 bp
* 84408 88430: contig of 4023 bp in length
* 88431 88530: gap of 100 bp
* 88531 94075: contig of 5545 bp in length
* 94076 94175: gap of 100 bp
* 94176 100299: contig of 6124 bp in length
* 100300 100399: gap of 100 bp
* 100400 107046: contig of 6647 bp in length
* 107047 107146: gap of 100 bp
* 107147 114514: contig of 7368 bp in length
* 114515 122515: contig of 100 bp
* 122516 122615: contig of 7901 bp in length
* 122616 131415: gap of 100 bp
* 131416 131515: contig of 8800 bp in length
* 131516 138482: contig of 6967 bp in length

* 138483 138582: gap of 100 bp
* 138583 144615: contig of 6033 bp in length
* 144616 144715: gap of 100 bp
* 144716 158405: contig of 13690 bp in length.
FEATURES
source
1. 158405
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-2P7"
/clone_lib="RPC1-11 Human Male BAC"
1. 1607
/note="assembly_fragment"
1708..3028
/note="assembly_fragment"
3129..4889
/note="assembly_fragment"
4990..6433
/note="assembly_fragment"
6534..8085
/note="assembly_fragment"
8186..9968
/note="assembly_fragment"
10069..11857
/note="assembly_fragment"
11958..14371
/note="assembly_fragment"
14472..16047
/note="assembly_fragment"
16148..18676
/note="assembly_fragment"
18777..21785
/note="assembly_fragment"
21886..24973
/note="assembly_fragment"
25074..27441
/note="assembly_fragment"
27542..30718
/note="assembly_fragment"
30819..33992
/note="assembly_fragment"
34093..36731
/note="assembly_fragment"
36832..39283
/note="assembly_fragment"
39384..41838
/note="assembly_fragment"
41939..44209
/note="assembly_fragment"
clone_end:SP6
vector_side:left
44310..47729
/note="assembly_fragment"
47830..50393
/note="assembly_fragment"
50494..56065
/note="assembly_fragment"
56166..61306
/note="assembly_fragment"
61407..65295
/note="assembly_fragment"
65396..71547

Query Match

27.0%; Score 232; DB 2; Length 158405;

Best Local Similarity 99.6%; Pred. No. 5.2e-114;

Matches 282; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 23 CGCCCTTCGGCCGGCCACCCACGCGCGTGTGTGGGGCGCTACCCGAGTCGCTTCGCAGC 82

Db 142537 CGCCCTTCGGCCGGCCACCCACGCGCGTGTGTGGGGCGCTACCCGAGTCGCTTCGCAGC 142596

Qy 83 ACGCCTGAGAGCGCCCAAGCGGAGGAGTGACGTGCGCCCGCGGACGACGACCC 142

```

Db 142597 ACGCGCTGAGAAGCGCGCAAGCGGAGGAGTGGACGTGCGCCGCGGGAACGGCAGCACC 142656
QY 143 AGCTCTACGTGGGCGTGGGAGAGAGCTGGGGTGGAGTGGTGGAGCTGCCGGCGC 202
Db 142657 AGCTCTACGTGGGCGTGGGAGAGAGCTGGGGTGGAGTGGTGGAGCTGCCGGCGC 142716
QY 203 ACAGAGACCTTCGCGAGCTGCGTCTTCGTGGAGGACGTGGCCGCTGGTGTGCGAGGAGACGG 262
Db 142717 ACAGAGACCTTCGCGAGCTGCGTCTTCGTGGAGGACGTGGCCGCTGGTGTGCGAGGAGACGG 142776
QY 263 CCTCTATCACCCAGCCCGGGCGCGAGCCGCGAGGAGGAGGAGT 305
Db 142777 CCTCTATCACCCAGCCCGGGCGCGAGCCGCGAGGAGGAGT 142819

RESULT 12
AL360219/c AL360219 168953 bp DNA linear PRI 27-FEB-2001
LOCUS Human DNA sequence from clone RP11-131L23 on chromosome 1, complete
DEFINITION sequence.
ACCESSION AL360219 AC073061
VERSION AL360219.18 GI:13169534
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Donnelly, S.
Direct Submission
Submitted (27-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On or before May 15, 2001 this sequence version replaced
GI:9958163, GI:12718084.
During sequence assembly data is compared from overlapping clones.
Where difference assemblies are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e. phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em.; EMBL; Sw.;
SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RP11-131L23 is from the library RPCT-11.1 constructed by the group
of Pieter de Jong. For further details see
http://pba.ce3.6
http://www.chori.org/bacpac/home.htm
IMPORTANT: This sequence is not the entire insert of clone
RP11-131L23 It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true right end of clone RP11-131L23 is at 168953 in this
sequence. The true left end of clone RP4-611J7 is at 37497 in this
sequence. The true left end of clone RP11-290M5 is at 104198 in
this sequence. The true right end of clone RP4-621F18 is at 100 in
this sequence. The true right end of clone RP4-611J7 is at 140647
in this sequence.

FEATURES
Location/Qualifiers
1..168953
/organism="Homo sapiens"

```

```

/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-131L23"
/clone_lib="RPCT-11.1"
13830..13855
/notes="Sequence confirmed by AC073061 sequenced by WUGSC"
27601..27792
/notes="Sequence confirmed by AC019261 sequenced by WIER"
147415..147467
/notes="Sequence from overlapping clone BA290M5 (AL162256).
Assembly confirmed by restriction digest."
161180..161206
/notes="Sequence from overlapping clone BA290M5 (AL162256).
Assembly confirmed by restriction digest."

ORIGIN
QY 23 CGCGCTTCGCGCGGCGCCACCCACCGCTGGTGGCGGGCGCTACCCGAGTGGCTCTGCCAGC 82
Db 74904 CGCGCTTCGCGCGGCGCCACCCACCGCTGGTGGCGGGCGCTACCCGAGTGGCTCTGCCAGC 74845
QY 83 ACGCGCTGAGAAAGCGCCCAAGCGGCGAGGAGTGGACGTGGCCGCGCGGAAACGGCAGCACC 142
Db 74844 ACGCGCTGAGAAAGCGCCCAAGCGGCGAGGAGTGGACGTGGCCGCGCGGAAACGGCAGCACC 74785
QY 143 AGCTCTACGTGGGCGTGGTGGGCGAGCAGCTGGGCGTGCAGTGGTGGAGTGGCGGCGC 202
Db 74784 AGCTCTACGTGGGCGTGGTGGGCGAGCAGCTGGGCGTGCAGTGGTGGAGTGGCGGCGC 74725
QY 203 ACGAGAGCCTTCGAGACTGCTCTTCGTGGAGGAGCTGGCCGCTGGTGGCAGGAGACGG 262
Db 74724 ACGAGAGCCTTCGAGACTGCTCTTCGTGGAGGAGCTGGCCGCTGGTGGCAGGAGACGG 74665
QY 263 CCTCTATCACCCAGCCCGGGCGCGAGCCGCGAGGAGGAGT 305
Db 74664 CCTCTATCACCCAGCCCGGGCGCGAGCCGCGAGGAGGAGT 74622

RESULT 13
HSA123M24/c HSA123M24 83946 bp DNA linear PRI 04-MAR-2003
LOCUS Human DNA sequence from clone RP4-621F18 on chromosome 1p11.4-21.3,
DEFINITION complete sequence.
ACCESSION AL078459
VERSION AL078459.8 GI:5791502
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 83946)
Whitehead, S.
Direct Submission
Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Aug 27, 1999 this sequence version replaced GI:5596949.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep -----
Genome Center

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 15:45:50 ; Search time 405 Seconds
(without alignments)
8999.886 Million cell updates/sec

Title: US-09-889-733B-1

Perfect score: 858

Sequence: 1 atggcggcctcgccacc.....acaagaagtagactcctga 858

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 337863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002s.*

7: Geneseqn2003as.*

8: Geneseqn2003bs.*

9: Geneseqn2003cs.*

10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	858	100.0	858	3	Aa47655 Dimethyla
2	851.6	99.3	858	9	Ade71465 Human Thr
3	851.6	99.3	1633	4	Aa402907 Human she
4	851.6	99.3	1633	6	Abq88102 Human ost
5	851.6	99.3	3932	7	Abz35939 Human sec
6	851.6	99.3	4208	6	Aa141494 Drug meta
7	850	99.1	858	9	Ade71463 Human Met
8	715.6	83.4	3008	9	Abz53014 Primary r
9	705.4	82.2	3008	9	Abz79843 Rat N-G,N
10	705.4	82.2	3778	9	Ades3981 Human pro
11	630.4	73.5	3651	3	Aa16077 Human pro
12	218.8	25.5	444	6	Abn89945 Mouse clo
13	216.2	25.2	1376	2	Aa252967 Human pro
14	214.6	25.0	858	3	Aa47656 Dimethyla
15	214.6	25.0	1228	3	Aa400677 Human Hyd
16	207.6	24.2	845	3	Aa99930 Human nel
17	202.6	23.6	862	6	Abq47888 Oligonucl
18	202.6	23.6	862	6	Abq47889 Oligonucl
19	202.6	23.6	898	6	Abq41319 Oligonucl
20	202.6	23.6	898	6	Abq41318 Oligonucl
21	197.8	23.1	770	6	Abz76755 Frog embr
22	167.8	19.6	862	6	Abq47896 Oligonucl
23	167.8	19.6	862	6	Abq47887 Oligonucl

ALIGNMENTS

RESULT 1

AAA47655

ID AAA47655 standard; cDNA; 858 BP.

XX AAA47655;

AC AAA47655;

DT 08-NOV-2000 (first entry)

DE Dimethylarginine dimethylaminohydrolase (DDAH1) coding sequence.

XX Dimethylarginine dimethylaminohydrolase; DDAH; DDAH1; DDAH2;

KW arginine deaminase; hyperlipidemia; renal failure; hypertension;

KW restenosis; atherosclerosis; schizophrenia; multiple sclerosis; cancer;

KW ischemia reperfusion injury; septic shock; multi organ failure;

KW arthritis; skin disorders; inflammatory cardiac disease; migraine;

XX infection; ds.

OS Homo sapiens.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

susceptibility to a cardiovascular disease and diabetes in an individual by determining the DDAH 1 genotype; a method of treating diabetes or vascular complications of diabetes by enhancing nitric oxide availability, production or concentration; a method of targeting treatment of cardiovascular disease and diabetes in a hypertensive patient by determining the DDAH 1 genotype and treating them with a drug that affects nitric oxide availability, production or metabolism; and a transgenic animal which carries a human variant DDAH 1 nucleic acid sequence. The nucleic acid molecules and polypeptides are useful for treating cardiovascular disease (e.g., coronary heart disease, cerebrovascular disease, and hypertension), and diabetes and its vascular complications. The methods are useful for determining whether a patient will benefit from treatment with a drug which affects nitric oxide availability, production or metabolism; a drug which reduces ADMA availability or concentration; or an agent which elevates DDAH availability or concentration (such as DDAH agonist). The methods are also useful for determining whether a patient will be at risk of adverse effects if DDAH antagonists are administered. The present sequence represents a nucleic acid encoding the "wild-type" (Thr87) DDAH 1.

```
Query Match          99.3%;   Score 851.6;   DB 9;   Length 858;
Best Local Similarity 99.5%;   Pred. No. 4.3e-184;
Matches 854: Conservative 0; Mismatches 4; Indels 0; Gaps 0
```

Qy	1	ATGCGCGGCTCGGCACACCCCTTCGCCTTCGGCCGGGCCACCCACGCGCTGTGTGCGGGCG	60
Db	1	ATGCGCGGCTCGGCACACCCCGCGCCTTCGGCCGGGCCACCCACGCGCTGTGTGCGGGCG	60
Qy	61	CTACCCGAGTCGCTCTGCGACGACGCGCTGAGAAGCCCAAGGGCGAGAGGTGACGCTC	120
Db	61	CTACCCGAGTCGCTCTGGCCAGCA CGGCTGAGAAAGCCCAAGGGCGAGAGGTGACGCTC	120
Qy	121	GCCCGCGGAAACGGCAGCA CCAAGCTCTA CTTGGGCGTGTGGCAGCAAGCTGTGGGCTGTG	180
Db	121	GCCCGCGGAAACGGCAGCA CCAAGCTCTTACCTGTGGCGCTGTGGCAGCAAGCTGTGGGCTGTG	180
Qy	181	CAGGTGTTGAGCTCGCGCCGACGAGAGCCTTCGGACTGCGCTTCCTGTGAGGACGCTG	240
Db	181	CAGGTGTTGAGCTCGCGCCGACGAGAGCCTTCGGACTGCGCTTCCTGTGTGAGGACGCTG	240
Qy	241	GCGTGTGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGGCGCGAGCCGAGGAAG	300
Db	241	GCGTGTGTGCGAGGAGACGGCCCTCATCACCCGACCCGGGGCGCGAGCCGAGGAAG	300
Qy	301	GAGTTGACATCATGAAAGAGCATTAGAAAACTTCAGCTCAATATAGTAGAGTGAAT	360
Db	301	GAGTTGACATCATGAAAGAGCATTAGAAAACTTCAGCTCAATATAGTAGAGTGAAT	360
Qy	361	GATGAAAAATGCAACTTTAGATGGCGGAGATGTTTTATTTCACGCGCAGAGAAATTTTTGTG	420
Db	361	GATGAAAAATGCAACTTTAGATGGCGGAGATGTTTTATTTCACGCGCAGAGAAATTTTTGTG	420
Qy	421	GGCCTTCCAAAAGGACAAATCAACGAGGTGCTGAATCTTGCTGNTACTTTTAAGAC	480
Db	421	GGCCTTCCAAAAGGACAAATCAACGAGGTGCTGAATCTTGCTGNTACTTTTAAGAC	480
Qy	481	TATGCACTCTCCACAGTCGACGTGGCAGATGGGTTCGATTTGAAGAGTTTCTCGACCATG	540
Db	481	TATGCACTCTCCACAGTGCAGTGGCAGATGGGTTCGATTTGAAGAGTTTCTCGACCATG	540
Qy	541	GCTGGCCCTAAACCTGATGCCAATTTGGGTCTTAGTGAATCTTCGACGAAGAGCCCTTTAAGTC	600
Db	541	GCTGGCCCTAAACCTGATGCCAATTTGGGTCTTAGTGAATCTTCGACGAAGAGCCCTTTAAGTC	600
Qy	601	ATGCAACAGATGAGTGACCAACCGCTACGACAAACTCACTGTGCTGATGACATAGCAGCA	660
Db	601	ATGCAACAGATGAGTGACCAACCGCTACGACAAACTCACTGTGCTGATGACATAGCAGCA	660
Qy	661	AACTGTATATATCTAAATATCCCCAACAAAGGCAACGCTTTGTGTGCAACCGAACCCCGAA	720
Db	661	AACTGTATATATCTAAATATCCCCAACAAAGGCAACGCTTTGTGTGCAACCGAACCCCGAA	720

QY	721	GAGTATCAGAAAGTGC	AAAGGTTTATGAGAA	CTGAAGACCATATGCTGATCCCGGTG	7801
DB	721	GAGTATCAGAAAGTGC	AAAGGTTTATGAGAA	CTGAAGACCATATGCTGATCCCGGTG	7801
QY	781	AGCATGCTGAAC	TGGAAAAAGTGGATGGGCTGCTACCTGCTGCTCAGT	TTTAAATTAAC	8401
DB	781	AGCATGCTGAAC	TGGAAAAAGTGGATGGGCTGCTACCTGCTGCTCAGT	TTTAAATTAAC	8401
QY	841	AAAGAGGTAGACT	CTCTGA	858	
DB	841	AAAGAGGTAGACT	CTCTGA	858	

RESULT 3
AAH02907
ID AAH02907 standard: DNA: 1633 BP.

XX	AAH02907;	
AC	15-JUN-2001 (first entry)	
XX	Human shear stress-response coding sequence SEQ ID NO: 67.	
XX	Human; shear stress-response protein; vascular disease; arteriosclerosis;	
DE	ds.	
DE	Homo sapiens.	
XX	WO200125427-A1.	
XX	12-APR-2001.	
XX	02-OCT-2000; 2000WO-JP006840.	
XX	01-OCT-1999; 99JP-00280976.	
XX	(KYOW) KYOWA HAKKO KOGYO KK.	
XX	(NOJI/) NOJIMA H.	
XX	Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;	
PI	Kuga T, Sekine S, Nakamura Y, Sugano S;	
PI	WPI; 2001-266308/27.	
DR	P-PSDB: AAB90784.	

PT DNA sequences, proteins encoded by them and antibodies against them
PT useful in diagnosis and treatment of vascular disease caused by
PT arteriosclerosis.

XX
PS Claim 20: Page 404-407: 678pp: Japanese:

xx The present invention provides the protein and coding sequences of a
cc number of human shear stress response proteins. These are useful in the
cc diagnosis, treatment and screening of vascular diseases caused by
cc arteriosclerosis, including heart failure, post-PICA stenosis and
cc hypertension

Sequence 1633 BP: 400 A: 406 C: 432 G: 393 T: 0 U: 2 Other:

Query Match	99.3%	Score 851.6;	DB 4;	Length 1633;
Best Local Similarity	99.5%;	Pred. No. 5.1e-194;		
Matches 854:	Conservative	0;	Mismatches 4;	Indels 0;
Gaps	0;			

Qy	1	ATGGCGCGCTCGGGCCACCCCTTCGCGCTTGGCGCGGGCGACCCACGCGCTGGTGGCGGGC	60
Db	323	ATGGCGCGGGCTCGGGCCACCCCGCGCGCTTGGCGCGGGCCACCCACGCGCTGGTGGCGGGCG	382
Qy	61	CTACCCGAGTCGCTCTGCCACACGCGCTGAGAAAGCCCAAGGGCGAGGAGGTGGACGTC	120
Db	383	CTACCCGAGTCGCTCTGGCGCAGCAGCGCTGAGAAAGCCCAAGGGCGAGGAGGTGGACGTC	442
Qy	121	GGCCCGCGGGAAACGGCAGCACAGCTCTTACGTGGGGCGCTGTGGCAGCAACAGCTGGGGCGTG	180


```
Db 443 GCCCGCGGAAACGGCAGACACAGCTCTACGTGGCGTCTGGGAGCAAGCTGGGGTG 502
QY 181 CAGGTGGTGGAGCTGCGCGCCACAGAGCCTTCCGGACTGCTTCTGTTGAGGACGTG 240
Db 503 CAGGTGGTGGAGCTGCGCGCCACAGAGCCTTCCGGACTGCTTCTGTTGAGGACGTG 562
QY 241 GCGGTGGTGGAGGAGACGGCCCTCATCACCGACCGCGGGCCCGAGCCGGAGGAAG 300
Db 563 GCGGTGGTGGAGGAGACGGCCCTCATCACCGACCGCGGGCCCGAGCCGGAGGAAG 622
QY 301 GAGGTGGATGATGATGAAAGACATTTAGAAAACTTACGCTCAATATAGTAGAGATGAAA 360
Db 623 GAGGTGGATGATGATGAAAGACATTTAGAAAACTTACGCTCAATATAGTAGAGATGAAA 682
QY 361 GATGAAATGCAACTTTAGATGGCGAGATGTTTTATTACAGGCAGAGAAATTTTTGTG 420
Db 683 GATGAAATGCAACTTTAGATGGCGAGATGTTTTATTACAGGCAGAGAAATTTTTGTG 742
QY 421 GCGCTTTCCAAAGACAAATCAACGAGGTGCTGAATCTTGGCTGATACATTTAAAGAC 480
Db 743 GCGCTTTCCAAAGACAAATCAACGAGGTGCTGAATCTTGGCTGATACATTTAAAGAC 802
QY 481 TATGCACTTCCACAGTGCAGTGGCAGATGGGTTGCAATTTGAAGAGTTTCTGCAGCATG 540
Db 803 TATGCACTTCCACAGTGCAGTGGCAGATGGGTTGCAATTTGAAGAGTTTCTGCAGCATG 862
QY 541 GCTGGCCCTAACCTGATCCCAATTTGGTCTAGTGAATCTGCACAGAGCCCTTAAGATC 600
Db 863 GCTGGCCCTAACCTGATCCCAATTTGGTCTAGTGAATCTGCACAGAGCCCTTAAGATC 922
QY 601 ATGCAACAGATGATGACACCGCTACGACAAACTCACTGTGCTGTGATGACATAGCAGCA 660
Db 923 ATGCAACAGATGATGACACCGCTACGACAAACTCACTGTGCTGTGATGACATAGCAGCA 982
QY 661 AACTGTATATATCAATATCCCAAAAGGGCAGCTTCTGCTCCACCGAACCCCGGAA 720
Db 983 AACTGTATATATCAATATCCCAAAAGGGCAGCTTCTGCTCCACCGAACCCCGGAA 1042
QY 721 GAGTATCCAGAAAGTCAAAAGTTTATGAAAGCTGAAGACCATATGCTGATCCCGTG 780
Db 1043 GAGTATCCAGAAAGTCAAAAGTTTATGAAAGCTGAAGACCATATGCTGATCCCGTG 1102
QY 781 AGCATGCTGAAGTGAAGAGTGGTGGCTGCTCACTGTGCTGTGATGATTTAATTAAAC 840
Db 1103 AGCATGCTGAAGTGAAGAGTGGTGGCTGCTCACTGTGCTGTGATGATTTAATTAAAC 1162
QY 841 AAGAGGTAGACTTCCTGA 858
Db 1163 AAGAAAGTAGACTTCCTGA 1180
```

```
RESULT 4
ID ABQ88102
XX ABQ88102 standard; cDNA; 1633 BP.
AC ABQ88102;
XX
DT 18-SEP-2002 (first entry)
XX
DE Human osteoblast differentiation related cDNA SEQ ID NO 9.
XX
KW Human; osteoblast; stem cell differentiation; bone tissue deposition;
XX osteoporosis; osteopathic; ss.
XX
OS Homo sapiens.
XX
PN WO20020301-A2.
XX
PD 27-JUN-2002.
XX
FF 18-DEC-2001; 2001WO-US048276.
XX
```

```
PR 18-DEC-2000; 2000US-0255882P.
XX 24-APR-2001; 2001US-0285691P.
PA (GENE-) GENE LOGIC INC.
PA (PROC ) PROCTER & GAMBLE CO.
XX
PI Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
PI Mertz L;
XX WPI; 2002-557663/59.
XX
PT Use of genes and their expression profiles associated with osteoblast
PT differentiation for screening modulators bone formation, for diagnosing
PT or treating e.g. osteoporosis, or as markers for the differentiation
PT process.
XX
PS Claim 1; SEQ ID NO 9; 78pp + Sequence Listing; English.
XX
CC The invention relates to genes and their expression profiles are used
CC for: (a) screening modulators of precursor stem cell differentiation into
CC osteoblasts or bone tissue deposition; (b) diagnosing abnormal
CC deposition of bone tissue, abnormal rate of osteoblast formation or
CC osteoporosis; or (c) treating or monitoring treatment of the conditions
CC cited in (b), or monitoring the progression of bone tissue deposition.
CC Specific conditions include postmenopausal osteoporosis, glucocorticoid
CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-
CC induced abnormalities in bone formation or bone loss, conditions that
CC involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
CC skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
CC or fibrous dysplasia. The present sequence is that of an osteoblast
CC differentiation associated cDNA marker of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1633 BP; 400 A; 406 C; 432 G; 393 T; 0 U; 2 Other;
```

```
Query Match 99.3%; Score 851.6; DB 6; Length 1633;
Best Local Similarity 99.5%; Pred. No. 5.1e-184;
Matches 854; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGCGCGGCTCGGCGCACCCCTCGGCTTCGGCGGGCCACCCACCCGCGGTGGCGGCG 60
Db 323 ATGCGCGGCTCGGCGCACCCCTCGGCTTCGGCGGGCCACCCACCCGCGGTGGCGGCG 382
QY 61 CTACCCGAGTCTGCTGCCAGCAGCGGCTGAGAAGCGCCAAAGGGCGAGGAGTGGACGTC 120
Db 383 CTACCCGAGTCTGCTGCCAGCAGCGGCTGAGAAGCGCCAAAGGGCGAGGAGTGGACGTC 442
QY 121 GCGCGCGCGAAGCGGAGACCGCCCTCATCACCGACCGCGGGCGGCGAGCCGAGGAAG 180
Db 443 GCGCGCGCGAAGCGGAGACCGCCCTCATCACCGACCGCGGGCGGCGAGCCGAGGAAG 502
QY 181 CAGGTGGTGGAGCTGCGCGCGGCGAGAGAGCTTTCGGAGCTTCGCTGGAGGACGTG 240
Db 503 CAGGTGGTGGAGCTGCGCGCGGCGAGAGAGCTTTCGGAGCTTCGCTGGAGGACGTG 562
QY 241 GCGGTGGTGGAGGAGACCGCCCTCATCACCGACCGCGGGCGGCGAGCCGAGGAAG 300
Db 563 GCGGTGGTGGAGGAGACCGCCCTCATCACCGACCGCGGGCGGCGAGCCGAGGAAG 622
QY 301 GAGGTTCACATGATGAAAGAGCATTTAGAAAACTTACGCTCAATATAGTAGAGATGAAA 360
Db 623 GAGGTTCACATGATGAAAGAGCATTTAGAAAACTTACGCTCAATATAGTAGAGATGAAA 682
QY 361 GATGAAATGCAACTTTAGATGGCGAGATGTTTTATTACAGGCAGAGAAATTTTTGTG 420
Db 683 GATGAAATGCAACTTTAGATGGCGAGATGTTTTATTACAGGCAGAGAAATTTTTGTG 742
QY 421 GCGCTTTCCAAAGACAAATCAACGAGGTGCTGAATCTTGGCTGATACATTTAAAGAC 480
Db 743 GCGCTTTCCAAAGACAAATCAACGAGGTGCTGAATCTTGGCTGATACATTTAAAGAC 802
```

QY 481 TATGAGTCTCCAGTGGCCAGTGGGAGTGGTTCATTTGAAGAGTTTCTGCAGCATG 540
DB 803 TATGAGTCTCCAGTGGCCAGTGGGAGTGGTTCATTTGAAGAGTTTCTGCAGCATG 862
QY 541 GCTGGGCTAACCTGATGCGCAATTTGGTCTAGTGAATCTGACAGAGCCCTTAAGATC 600
DB 863 GCTGGGCTAACCTGATGCGCAATTTGGTCTAGTGAATCTGACAGAGCCCTTAAGATC 922
QY 501 ATGCACAGATGATGACACCGCTACGACAACTCACTGCTGATGACATAGCAGCA 560
DB 923 ATGCACAGATGATGACACCGCTACGACAACTCACTGCTGATGACATAGCAGCA 982
QY 661 AACTGTATATATCTAAATATCCCAACAAAGGGGACGCTTGTGTCACCGAACCCCGGAA 720
DB 983 AACTGTATATCTAAATATCCCAACAAAGGGGACGCTTGTGTCACCGAACCCCGGAA 1042
QY 721 GAGTATCCAGAGTGCAGAGGTTTATGAGAACTGAGAGACCATATGCTGATCCCGGTG 780
DB 1043 GAGTATCCAGAGTGCAGAGGTTTATGAGAACTGAGAGACCATATGCTGATCCCGGTG 1102
QY 781 AGCATGCTGAACCTGGAAGAGGTTGATGGGCTGCTACCTGCTGCTCAAGTTTAAATTAAC 840
DB 1103 AGCATGCTGAACCTGGAAGAGGTTGATGGGCTGCTACCTGCTGCTCAAGTTTAAATTAAC 1162
QY 841 AAGAAGGTAGACTCTCGA 858
DB 1163 AAGAAGGTAGACTCTCGA 1180

RESULT 5
ID ABZ35939 standard; cdNA; 3932 BP.
XX AC ABZ35939;
XX DT 10-FEB-2003 (first entry)
XX DE Human secretory polynucleotide SPTM SEQ ID NO 103.
XX KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;
XX KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;
XX KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;
XX KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;
XX KW anti-inflammatory; immunosuppressive; neuroprotective; nontropic;
XX KW neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;
XX KW antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;
XX KW secretory polynucleotide; secretory protein; gene; ss.
XX OS Homo sapiens.
XX PN WO2002083876-A2.
XX PD 24-OCT-2002.
XX PF 27-MAR-2002; 2002WO-US009921.
XX PR 29-MAR-2001; 2001US-0280067P.
XX PR 29-MAR-2001; 2001US-0280068P.
XX PR 16-MAY-2001; 2001US-0291280P.
XX PR 17-MAY-2001; 2001US-0291829P.
XX PR 17-MAY-2001; 2001US-0291849P.
XX PR 19-JUN-2001; 2001US-0299428P.
XX PR 20-JUN-2001; 2001US-0299776P.
XX PR 20-JUN-2001; 2001US-030001P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Datto A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
XX PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
XX PI Daugherty SC, Lam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstein EH;
XX PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

DR MPI: 2003-075543/07.
DR P-PSDB; ABP75492.
XX New human secretory proteins and polynucleotides, useful for diagnosing, treating or preventing autoimmune/inflammatory disorders (e.g. AIDS), neurological disorders (e.g. Alzheimer's), or cell proliferations or cancers.
PS Claim 1; SEQ ID NO 103; 458pp + Sequence Listing; English.
XX The invention relates to a secretory polynucleotide (designated spcm) comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a naturally occurring polynucleotide sequence at least 90 % identical to the polynucleotide sequence, a polynucleotide complementary to them or an RNA equivalent of them. The polypeptide or polynucleotide are useful for treating, preventing or diagnosing a disease or condition associated with the expression of functional SPTM. These are particularly useful for diagnosing, treating or preventing autoimmune/inflammatory disorders (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's disease), neurological disorders (e.g. epilepsy, Huntington's disease, dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease, multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety, schizophrenia or amnesia), or cell proliferative disorders (e.g. psoriasis, polycythemia vera, or cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain, breast, cervix or prostate). Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 3932 BP; 1089 A; 883 C; 775 G; 1185 T; 0 U; 0 Other;
Query Match 99.3%; Score 851.6; DB 7; Length 3932;
Best Local Similarity 99.5%; Pred. NO. 6.5e-184;
Matches 854; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGGCGGCGCTCGGCCACCCCTCCGCTCGCGGGGACCCACCGCTGTGGCGGCG 60
DB 17 ATGGCGGCGCTCGGCCACCCCTCGCGGGGACCCACCGCTGTGGCGGCG 76
QY 61 CTACCCGAGTCTCTGCCAGCACCGCTGAGAGCGCAAGGGCGAGAGTGACGTC 120
DB 77 CTACCCGAGTCTCTGCCAGCACCGCTGAGAGCGCAAGGGCGAGAGTGACGTC 136
QY 121 GCCCGCGGAGCGACACCCAGCTCTACGTGGGCGTCTGGGCGAGCAAGCTGGGCGTG 180
DB 137 GCCCGCGGAGCGACACCCAGCTCTACGTGGGCGTCTGGGCGAGCAAGCTGGGCGTG 196
QY 181 CAGGTGGTGGAGTCTCGCGGCCGACGAGAGCTTCGCGACTGCTCTCGTGAGGACGTG 240
DB 197 CAGGTGGTGGAGTCTCGCGGCCGACGAGAGCTTCGCGACTGCTCTCGTGAGGACGTG 256
QY 241 GCCGTGGTGGAGGAGAGCGCCCTCATCCCGACCCCGGGGCCCGAGCGGAGAG 300
DB 257 GCCGTGGTGGAGGAGAGCGCCCTCATCCCGACCCCGGGGCCCGAGCGGAGAG 316
QY 301 GAGGTTGACATGATGAAAGAGCATTAGAAAACCTTCAGCTCAATATAGTAGAGTAA 360
DB 317 GAGGTTGACATGATGAAAGAGCATTAGAAAACCTTCAGCTCAATATAGTAGAGTAA 376
QY 361 GATGAAATGCAACTTTTAGATGCGGAGATGTTTATTACAGGCGAGAGAAATTTTGTG 420
DB 377 GATGAAATGCAACTTTTAGATGCGGAGATGTTTATTACAGGCGAGAGAAATTTTGTG 436
QY 421 GGCCTTTCCAAAGGACAAATCAAGAGGTGCTGAATCTTCGCTGATCTTTTAAGGAC 480
DB 437 GGCCTTTCCAAAGGACAAATCAAGAGGTGCTGAATCTTCGCTGATCTTTTAAGGAC 496
QY 481 TATGAGTCTCCAGTGGCCAGTGGGAGTGGTTCATTTGAAGAGTTTCTGCAGCATG 540
DB 497 TATGAGTCTCCAGTGGCCAGTGGGAGTGGTTCATTTGAAGAGTTTCTGCAGCATG 556
QY 541 GCTGGGCTAACCTGATGCGCAATTTGGTCTAGTGAATCTGACAGAGCCCTTAAGATC 600

QY 481 TATGAGTCTCCAGTGCAGTGCAGATGGTTCATTTGAAGAGTTTCTGCAGCATG 540
 DB 848 TATGAGTCTCCAGTGCAGTGCAGATGGTTCATTTGAAGAGTTTCTGCAGCATG 907
 QY 541 GCTGGGCTAACCTGATGCGAATTTGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATC 600
 DB 908 GCTGGGCTAACCTGATGCGAATTTGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATC 967
 QY 601 ATGCAACAGATGAGTGCACCGCTACGACAACTCACTGTGCTGATGACATAGCAGCA 660
 DB 968 ATGCAACAGATGAGTGCACCGCTACGACAACTCACTGTGCTGATGACATAGCAGCA 1027
 QY 661 AACTGTATATCTAATATATCCCAAGAGGCACTCTTGTGTCACCGAACCCCGGAA 720
 DB 1028 AACTGTATATCTAATATATCCCAAGAGGCACTCTTGTGTCACCGAACCCCGGAA 1087
 QY 721 GAGTATCCAGAACTGCAAGGTTTATGAGAACTGAGGACCATATGCTGATCCCGTG 780
 DB 1088 GAGTATCCAGAACTGCAAGGTTTATGAGAACTGAGGACCATATGCTGATCCCGTG 1147
 QY 781 AGCATGCTGAACCTGGAAGGTTGATGGGCTGCTCACCTGCTGCTCACTGCTGCTTAATTAAC 840
 DB 1148 AGCATGCTGAACCTGGAAGGTTGATGGGCTGCTCACCTGCTGCTCACTGCTGCTTAATTAAC 1207
 QY 841 AAGAAAGTACTCTCTGA 858
 DB 1208 AAGAAAGTACTCTCTGA 1225

RESULT 7
 ADE71463
 ID ADE71463 standard; DNA; 858 BP.
 AC ADE71463;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human Met87 variant dimethylarginine dimethylaminohydroxylase-1 DNA.
 KW Human; variant dimethylarginine dimethylaminohydroxylase-1; variant DDAH 1;
 KW chromosome 1; single nucleotide polymorphism; SNP; ADMA;
 KW asymmetrical dimethyl-arginine; nitric oxide; cardiovascular disease;
 KW coronary heart disease; cerebrovascular disease; hypertension; diabetes;
 KW susceptibility; genotyping; therapeutic targeting; antidiabetic;
 KW cardiant; cerebroprotective; hypotensive; gene therapy; Gene; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 1..858
 FT /*tag= a
 FT /product= "Met87 variant DDAH 1"
 FT replace(260,C)
 FT /*tag= b
 FT /standard name= "Single nucleotide polymorphism (SNP)"
 FT /note= "Results in Thr (C variant) or Met (T variant) at
 FT position 87 of the protein"
 XX
 PN WO2003089638-A1.
 XX
 PD 30-OCT-2003.
 XX
 PF 11-APR-2003; 2003WO-FI000274.
 XX
 PR 19-APR-2002; 2002US-00125456.
 XX
 PA (JURI-) JURILAB LTD OY.
 XX
 PI Valkonen V, Salonen JT, Pireskanen M, Tuomainen T, Laakso J;
 PI Laaksonen R;
 XX
 XX WPI; 2003-854121/79.

DR P-PSDB; ADE71464.
 XX
 PT New nucleic acid encoding a variant dimethylarginine
 PT dimethylaminohydroxylase-1 (DDAH-1) protein, useful for treating diabetes,
 PT and its vascular complications, e.g. coronary or cerebrovascular disease
 PT or hypertension.
 XX
 PS Claim 4; SEQ ID NO 1; 37pp; English.
 XX
 CC The invention relates to a variant dimethylarginine
 CC dimethylaminohydroxylase-1 (DDAH 1; ADE71464) and a nucleic acid encoding
 CC it (ADE71463). DDAs regulate the metabolism of ADMA (asymmetrical
 CC dimethyl-arginine), a naturally occurring nitric oxide synthase inhibitor)
 CC and nitric oxide, which are relevant in disorders such as cardiovascular
 CC disease and diabetes. The DDAH 1 gene is located on chromosome 1, which
 CC has been implicated in susceptibility to diseases such as familial
 CC combined hyperlipidaemia, premature coronary artery disease (CAD), non
 CC insulin-dependent (type 2) diabetes mellitus, and diastolic hypertension.
 CC The variant DDAH 1 of the invention has Met at position 87, rather than
 CC Thr, which is caused by a C to T change in exon 1 at a position hetero-
 CC corresponding to base 260 in ADE71463 and ADE71465. Individuals hetero-
 CC or homozygous for the Thr87Met allele have an increased risk of
 CC developing cardiovascular diseases and diabetes. The invention also
 CC relates to a DDAH 1 cDNA capture probe; a method and kit for diagnosing a
 CC susceptibility to a cardiovascular disease and diabetes in an individual
 CC by determining the DDAH 1 genotype; a method of treating diabetes or
 CC vascular complications of diabetes by enhancing nitric oxide
 CC availability, production or concentration; a method of targeting
 CC treatment of cardiovascular disease and diabetes in a hypertensive
 CC patient by determining the DDAH 1 genotype and treating them with a drug
 CC that affects nitric oxide availability, production or metabolism; and a
 CC transgenic animal which carries a human variant DDAH 1 nucleic acid
 CC sequence. The nucleic acid molecules and polypeptides are useful for
 CC treating cardiovascular disease (e.g., coronary heart disease,
 CC cerebrovascular disease, and hypertension), and diabetes and its vascular
 CC complications. The methods are useful for determining whether a patient
 CC will benefit from treatment with a drug which affects nitric oxide
 CC availability, production or metabolism; a drug which reduces ADMA
 CC availability or concentration; or an agent which elevates DDAs
 CC availability or concentration (such as DDAH agonist). The methods are
 CC also useful for determining whether a patient will be at risk of adverse
 CC effects if DDAH antagonists are administered. The present sequence
 CC represents a specifically claimed nucleic acid encoding the variant
 CC (Met87) DDAH 1.
 XX
 SQ Sequence 858 BP; 218 A; 212 C; 258 G; 170 T; 0 U; 0 Other;

Query Match 99.1%; Score 850; DB 9; Length 858;
 Best Local Similarity 99.4%; Pred. No. 9.9e-184;
 Matches 853; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGGCCGGCTCGGCCACCCCTCCGCTTCGGCCGGGCCACCCAGCCCTGCTGGGGGG 60
 DB 1 ATGGCCGGCTCGGCCACCCCTCCGCTTCGGCCGGGCCACCCAGCCCTGCTGGGGGG 60
 QY 61 CTACCCGAGTCTGTCAGCAGCAGCGCTGAGAGCGCCAGCGGAGGAGTGCACGTC 120
 DB 61 CTACCCGAGTCTGTCAGCAGCAGCGCTGAGAGCGCCAGCGGAGGAGTGCACGTC 120
 QY 121 GCCCGCGGGAACGCGCAGCAGCAGCTCTACGTCGGCGCTGCTGGCAGCAAGCTGGGGCTG 180
 DB 121 GCCCGCGGGAACGCGCAGCAGCAGCTCTACGTCGGCGCTGCTGGCAGCAAGCTGGGGCTG 180
 QY 181 CAGGTGGTGGAGCTGCCGGCCGACGAGAGCTTCGGACTGCGTCTTCGTGAGGACGTCG 240
 DB 181 CAGGTGGTGGAGCTGCCGGCCGACGAGAGCTTCGGACTGCGTCTTCGTGAGGACGTCG 240
 QY 241 GCCGTGGTGGAGGAGCGCCCTCATCCCGCCCGGGGCGCCGAGCGGAGGAG 300
 DB 241 GCCGTGGTGGAGGAGCGCCCTCATCCCGCCCGGGGCGCCGAGCGGAGGAG 300
 QY 301 GAGGTGACATGATGAAGAAGCATTAGAAAAAATTCAGCTCAATATAGTAGATGAA 360

481 TATGAGTCTCCACAGTGGCAGATGGGTGCAATTTGAAGATTTCTGCAGCATG 540
912 TACGCAATTTCCACAGTCCCGTGGCGGATCTTTGCAATTTAAAGATTTCTGCAGCATG 971
541 GCTGGGCTTACCTGATGCGCAATGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATC 600
972 GCTGGGCTTACCTGATGCGCAATGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATC 1031
601 ATGCAACAGATGATGACACCGCTAGCAGCAAACTCACTGCTGCTGATGACATAGCAGCA 660
1032 ATGCAACAGATGATGACACCGCTAGTGAAGCTTATGCAAGCTTACTGACGACGATGGCGGC 1091
661 AACTGTATATATCTAAATATCCCAACAAAGGCGACGTCTTGTGTCACCGAACCCCGGAA 720
1092 AACTGTATATATCTAAATATCCCAACAAAGGCGACGTCTTGTGTCACCGAACCCCGGAA 1151
721 GAGTATCCAGAAAGTCAAAAGGTTTATGAGAACTGAGAACCATATGCTGATCCCGGTG 780
1152 GAGTATCCAGAAAGTCAAAAGGTTTATGAGAACTGAGAACCATATGCTGATCCCGGTG 1211
781 AGCATGCTCAACTGGAAAGGTTGATGGGCTGCTCACCTGCTGCTCAGTTTAAATTAAC 840
1212 AGCAATCTGATGATGAGAAAGGTTGAGAACTGAGAACCATATGCTGATCCCGGTG 1271
841 AAGAAGTACTCTCTGA 858
1272 AAGAAGTACTCTTGA 1289

RESULT 9

ADB79843
ID ADB79843 standard; DNA; 3008 BP.
XX AC ADB79843;
XX DT 04-DEC-2003 (first entry)
XX DE Rat N-G,N-G-dimethylarginine dimethylaminohydrolase DNA, SEQ ID 83.
XX KW Analgesic; pain; streptozocin-induced diabetes; rat; gene; ds.
XX OS Rattus norvegicus.
XX PN EPI279744-A2.
XX PD 29-JAN-2003.
XX PF 26-JUL-2002; 2002EP-00255249.
XX PR 27-JUL-2001; 2001GB-00018354.
XX PR 07-FEB-2002; 2002GB-00002910.
XX PA (WARN) WARNER LAMBERT CO.
XX PI Brooksbank RA, Dixon AK, Lee K, Pinnock RD;
XX DR WPI; 2003-395407/38.
XX DR P-PSDB; ADB79842.

Use of isolated gene sequences and encoded polypeptides that are upregulated in the spinal cord in response to streptozocin-induced diabetes for screening compounds for the treatment of pain, or for diagnosing pain.

Claim 1; Page 155; 334pp; English.

The present invention relates to nucleotide sequences which are useful in the screening of compounds for the treatment of pain, or for the diagnosis of pain. The nucleotide sequences are up-regulated in the spinal cord in response to streptozocin-induced diabetes. The present sequence is one such nucleotide sequence.

Sequence 3008 BP; 766 A; 842 C; 658 G; 740 T; 0 U; 2 Other;
Query Match 83.4%; Score 715.6; DB 9; Length 3008;
Best Local Similarity 89.6%; Pred. No. 5.5e-153;
Matches 769; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
QY 1 ATGGCCGCGCTCGGCCACCCCTCCGCTTCGGCGCGGCCACCCACGCGTGTGGCGCG 60
DB 432 ATGGCCGCGCTCAGCCACCCCTCGTCTTCGGCGCGGCCACCCACGCGTGTGGCGCG 491
QY 61 CTATCCGAGTCTGCTGCGCAGCAGCGCTGAGAGCGCCAGGGCGGAGGAGGTGACCTC 120
DB 492 CGCCCGAGTCTGCTGCGCCAGCGCTGAGGCGCTCCAGGGCGGAGGAGGTGATTC 551
QY 121 GCCCGCGCGGAACGGCAGCAGCAGCTCTACGTGGCGCGTGTGGCGCAGCAAGCTGGGCG 180
DB 552 GCTCGCGCTGAGCGCCAGCAGCAGCTCTACGTGGCGCGTGTGGCGCAGCAAGCTGGGCG 611
QY 181 CAGGTGTGGAGCTCCCGCGCAGCAGAGCCCTTCGGAGCTGCTCTGCTGGAGACGTG 240
DB 612 CAGGTGTGGAGCTCCCGCGCAGCAGAGCCCTTCGGAGCTGCTCTGCTGGAGACGTG 671
QY 241 GCCGTGGTGTGGAGGAGCAGCGCCCTCATCACCGCGCGCGCGCGCGCGAGCGGAGGAG 300
DB 672 GCCGTGGTGTGGAGGAGCAGCGCCCTCATCACCGCGCGCGCGCGCGCGAGCGGAGG 731
QY 301 GAGGTGTGATGAGAAAGAGCAATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
DB 732 GAGGTGTGATGAGAAAGAGCCTTTGGAAAACTTCAGCTCAATATAGTAGAGATGAAA 791
QY 361 GATGAAATGCAACTTTAGATGGCGGAGATGTTTTATTTCAGGCGAGAGAAATTTTCTG 420
DB 792 GATGAAATGCAACTTTAGATGGCGGAGATGTTTTATTTCAGGCGAGAGAAATTTTCTG 851
QY 421 GGCCTTTCCAAAGGACAAATCAACAGAGTCTGAGATCTTTGGCTGATATCTTTAAGAC 480
DB 852 GGCCTTTCCAAAGGACAAATCAACAGAGTCTGAGATCTTTGGCTGATATCTTTAAGAC 911
QY 481 TATGAGTCTCCACAGTCCAGTGGCAGATGGGTGCAATTTGAAGATTTCTGCAGCATG 540
DB 912 TACGCAATTTCCACAGTCCCGTGGCGGATCTTTGCAATTTAAAGATTTCTGCAGCATG 971
QY 541 GCTGGGCTTACCTGATGCGCAATTTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATC 600
DB 972 GCTGGGCTTACCTGATGCGCAATTTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATC 1031
QY 601 ATGCAACAGATGATGACACCGCTAGCAGCAAACTCACTGCTGCTGATGACATAGCAGCA 660
DB 1032 ATGCAACAGATGATGACACCGCTAGCAGCAAACTCACTGCTGCTGATGACATAGCAGCA 1091
QY 661 AACTGTATATATCTAAATATCCCAACAAAGGCGACGTCTTGTGTCACCGAACCCCGGAA 720
DB 1092 AACTGTATATATCTAAATATCCCAACAAAGGCGACGTCTTGTGTCACCGAACCCCGGAA 1151
QY 721 GAGTATCCAGAAAGTCAAAAGGTTTATGAGAACTGAGAACCATATGCTGATCCCGGTG 780
DB 1152 GAGTATCCAGAAAGTCAAAAGGTTTATGAGAACTGAGAACCATATGCTGATCCCGGTG 1211
QY 781 AGCATGCTCAACTGGAAAGGTTGATGGGCTGCTCACCTGCTGCTCAGTTTAAATTAAC 840
DB 1212 AGCAATCTGATGATGAGAAAGGTTGAGAACTGAGAACCATATGCTGATCCCGGTG 1271
QY 841 AAGAAGTACTCTCTGA 858
DB 1272 AAGAAGTACTCTTGA 1289

RESULT 10

ADB53981
ID ADB53981 standard; cDNA; 3778 BP.
XX AC ADB53981;
XX AC ADB53981;

DT 29-JAN-2004 (first entry)
XX Human prostate cancer cDNA #328.
DE
XX Human; prostate cancer; ss; cDNA combination; differential expression;
KW gene.
KW
XX Homo sapiens.
OS
XX US2003190640-A1.
FN
XX 09-OCT-2003.
PD
XX 29-MAY-2002; 2002US-00252157.
PF
XX 31-MAY-2001; 2001US-0295048P.
PR
XX (FARI//) FARIS M.
PA (PEAR//) PEARSON C I.
XX
XX Faris M, Pearson CI;
PI
XX WPI; 2003-831619/77.
DR
XX New combination comprising cDNAs that are differentially expressed in
PT prostate cancer, useful for diagnosing, treating or monitoring the
PT progression of treatment of prostate cancer.
XX
XX Claim 1; SEQ ID NO 328; 42pp; English.
XX
XX The invention relates to a combination comprising a number of cDNAs
CC expressed in prostate cancer. The invention also relates to a method for
CC detecting differential expression of one or more cDNAs in a sample
CC containing nucleic acids by hybridising a substrate with the nucleic
CC acids, thus forming one or more hybridisation complexes, detecting
CC hybridisation complex formation and comparing the complexes formed with
CC standard complexes, where differences between the standard and the sample
CC complex formation indicate differential expression of cDNAs in the
CC sample. The differential expression is diagnostic of prostate cancer. The
CC invention also relates to proteins and antibodies related to the cDNAs.
CC The combination is useful for diagnosing, treating or monitoring the
CC progression of treatment of prostate cancer. The antibodies are useful
CC for detecting prostate cancer. This sequence represents a human prostate
CC cancer cDNA of the invention.
XX
SQ Sequence 3778 BP; 1064 A; 821 C; 719 G; 1150 T; 0 U; 24 Other;
Query Match 82.2%; Score 705.4; DB 9; Length 3778;
Best Local Similarity 99.7%; Pred. No. 1.2e-150;
Matches 717; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
141 CCAGCTCTACGTGGCGTCTGGCGAGCAAGCTGGGCTGCAGGTGGTGGAGCTGCCGGC 200
1 CCAGCTCTACGTGGCGTCTGGCGAGCAAGCTGGGCTGCAGGTGGTGGAGCTGCCGGC 60
201 CGACGAGAGCTTCGGAGCTGCTTCTGGAGGACGTGGCGTGGTGGTGGAGGAGAC 260
61 CGACGAGAGCTTCGGAGCTGCTTCTGGAGGACGTGGCGTGGTGGTGGAGGAGAC 120
261 GGCCCTCTACCTACCGACCGCGGCGCGCGAGCGGAGGAGGTTGACATGATGAAGA 320
121 GGCCCTCTACCTACCGACCGCGGCGCGCGAGCGGAGGAGGTTGACATGATGAAGA 180
321 AGCATTAGAAAACCTTACGCTCAATATAGTAGATGAAGATGAATGAATGAATGA 380
181 AGCATTAGAAAACCTTACGCTCAATATAGTAGATGAAGATGAATGAATGAATGA 240
381 TGGCGGAGATGTTTATTACAGGCGAGAGAAATTTTCTGGGCTTTCCAAAAGCAAA 440
241 TGGCGGAGATGTTTATTACAGGCGAGAGAAATTTTCTGGGCTTTCCAAAAGCAAA 300
441 T-CACGAGGTGCTGAATCTTGGCTGATACCTTTTAAGGACATGACGCTCCACAGTGC 499

Db 301 TCCACGAGGTGCTGAATCTTGGCTGATACCTTTAAGGACTATGACGTCTCCACAGTGC 360
QY 500 CAGTGGCAGATGGTTCATTTGAAGAGTTTCTGCAGATGGCTGGGCTAAACCTGATGC 559
Db 361 CAGTGGCAGATGGTTCATTTGAAGAGTTTCTGCAGATGGCTGGGCTAAACCTGATGC 420
QY 560 CAATTGGGTCTAGTGAATCTGCACAGAGGCGCTTAAAGATCATGCAACAGATGAGTGACC 619
Db 421 CAATTGGGTCTAGTGAATCTGCACAGAGGCGCTTAAAGATCATGCAACAGATGAGTGACC 480
QY 620 ACCGCTACGACAAACTCACTGTGCTGATGACATAGCAGCAAACTGTATATCTAAATA 679
Db 481 ACCGCTACGACAAACTCACTGTGCTGATGACATAGCAGCAAACTGTATATCTAAATA 540
QY 680 TCCCCACAAAGGGCAGCTTCTGTCACCCAGAACCCCGGAGAGTATCCAGAAAGTGCA 739
Db 541 TCCCCACAAAGGGCAGCTTCTGTCACCCAGAACCCCGGAGAGTATCCAGAAAGTGCA 600
QY 740 AGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCGTGAGCATGTCTGAACCTGAAA 799
Db 601 AGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCGTGAGCATGTCTGAACCTGAAA 660
QY 800 AGGTGATGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 858
Db 661 AGGTGATGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 719
RESULT 11
AAF16077
ID AAF16077 standard; cDNA; 3651 BP.
XX
AC AAF16077;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:512.
XX
KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytostatic; cardiac; immunomodulatory; muscular;
KW vulnary; gastrointestinal; nephrotropic; anti-infective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease; ss.
XX
OS Homo sapiens.
XX
XX WO200055174-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000MO-US005988.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM;
XX
XX WPI; 2000-587513/55.
DR
DR P-PSDB; AAB56874.
XX
PT Prostate cancer associated gene sequences, referred to as prostate cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as prostate cancer.
XX
XX Claim 1; Page 976-977; 2338pp; English.
XX
XX AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardiac, immunomodulatory, muscular, vulnary, gastrointestinal,
CC nephrotropic, anti-infective, gynaecological and antibacterial activities,

CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 3651 BP; 1016 A; 801 C; 684 G; 1141 T; 0 U; 9 Other;
Query Match 73.5%; Score 630.4; DB 3; Length 3651;
Best Local Similarity 99.7%; Pred. No. 1.4e-133;
Matches 642; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 215 CGGACTGCGTTCCTGCGAGAGAGTGGCGGTGTGTCGAGGAGCGGCCCTCATCACCC 274
Db 2 CGGACTGCGTTCCTGCGAGAGAGTGGCGGTGTGTCGAGGAGCGGCCCTCATCACCC 61
QY 275 GACCCGGCGCGCCGAGCGGAGGAGGAGTTCACATGATGAAGAAGCATTAGAAAAC 334
Db 62 GACCCGGCGCGCCGAGCGGAGGAGGAGTTCACATGATGAAGAAGCATTAGAAAAC 121
QY 335 TTGAGCTCAATATAGTAGAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAG 394
Db 122 TTGAGCTCAATATAGTAGAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAG 181
QY 395 TATTCAAGCAGCAGAGATTTTGTGGCGCTTTCCAAAGCAGCAATCAACGAGGTGCTG 454
Db 182 TATTCAAGCAGCAGAGATTTTGTGGCGCTTTCCAAAGCAGCAATCAACGAGGTGCTG 241
QY 455 AATCTTGGCTGTATCTTTTAAAGCATATGACATGTCACAGTGCAGTGGCGAGTGGGT 514
Db 242 AATCTTGGCTGTATCTTTTAAAGCATATGACATGTCACAGTGCAGTGGCGAGTGGGT 301
QY 515 TGCAATTTGAAGAGTTTCTGCGAGATGCTGGCGCTAACCTGATCGCAATTGGGTCTAGTG 574
Db 302 TGCAATTTGAAGAGTTTCTGCGAGATGCTGGCGCTAACCTGATCGCAATTGGGTCTAGTG 361
QY 575 AATCTGACAGAGAGCGCTTTAAGATCATGCAACAGATGATGACACCGCTACGACAAAC 634
Db 362 AATCTGACAGAGAGCGCTTTAAGATCATGCAACAGATGATGACACCGCTACGACAAAC 421
QY 635 TCAGTGTGCTGTATGATAGCAGCAAGCACTGTATATCTTAAATATCCCAACAGAGGCG 694
Db 422 TCAGTGTGCTGTATGATAGCAGCAAGCACTGTATATCTTAAATATCCCAACAGAGGCG 481
QY 695 ACCTCTTGTGTCACCGAACCCCGAGAGATATCCAGAAAGTGCAAAAGTTTATGAGAAAC 754
Db 482 ACCTCTTGTGTCACCGAACCCCGAGAGATATCCAGAAAGTGCAAAAGTTTATGAGAAAC 541
QY 755 TGAAGGACCATATGCTGATCCCGTGAGCATGCTGAACTGGAAAAGGTGGATGGGCTGC 814
Db 542 TGAAGGACCATATGCTGATCCCGTGAGCATGCTGAACTGGAAAAGGTGGATGGGCTGC 601
QY 815 TCACCTGCTGCTCAGTTTAAATTAACAGAGAGGTAGACTCCTGA 858
Db 602 TCACCTGCTG-TCAGTTTAAATTAACAGAGAGGTAGACTCCTGA 644

RESULT 12
ABN89945
ID ABN89945 standard; cDNA; 444 BP.
XX AC ABN89945;
XX DT 16-AUG-2002 (first entry)
XX DE Mouse clone DST IMX3_103 cDNA.
XX KW Mouse; antiinflammatory; gene therapy; ileitis; DST; ss; TOGA;
XX digital sequence tag; total gene expression analysis.

XX Mus musculus.
OS WO200231114-A2.
XX
XX 18-APR-2002.
XX
XX 11-OCT-2001; 2001WO-US032091.
XX
XX 11-OCT-2000; 2000US-0239483P.
XX
XX (DIGI-) DIGITAL GENE TECHNOLOGIES INC.
XX
XX Viney JL, Sims JE, Dubose RF, Baum PR, Hasel KW, Hilbush BS;
XX WPI; 2002-426279/45.
XX
XX New isolated nucleic acid molecules that are associated with ileitis, for
XX preventing, treating, modulating and diagnosing ileitis in a mammalian
XX subject.
XX
XX Claim 1; Page 212-213; 273pp; English.
XX
XX The invention relates to a novel isolated nucleic acid molecule
XX comprising a polynucleotide having one of 90 polynucleotide sequences,
XX given in the specification. The polynucleotides of the invention have
XX antiinflammatory activity, and may have a use in gene therapy. The
XX polynucleotide or a polypeptide encoded by it is used for preventing,
XX treating, modulating or ameliorating a medical condition such as ileitis.
XX The polypeptide or polynucleotide is also useful for manufacturing a
XX medicament for treating ileitis. The sequence represents a cDNA digital
XX sequence tag obtained from a mouse clone by the TOGA (total gene
XX expression analysis) method
XX
XX SQ Sequence 444 BP; 119 A; 126 C; 101 G; 98 T; 0 U; 0 Other;
Query Match 25.5%; Score 218.8; DB 6; Length 444;
Best Local Similarity 82.7%; Pred. No. 4e-40;
Matches 263; Conservative 0; Mismatches 52; Indels 3; Gaps 1;
QY 544 GGGCTTAACCTGATCGCAATGCGTCTAGTGAATCTGCAGAGAGCGCCCTTAAGATCATG 603
Db 2 GGACCCAACTGATGCAATAGGTGCTAGGTCAGCGAATCTGCAGAGAGCGCCCTCAAGATCATG 61
QY 604 CAACAGATGAGTGACCCCGC---TACGACAAACTCACTGTGCTGTATGACATAGCAGCA 660
Db 62 CAACAGATGAGTGACCCCGCCTGTTAATGCGAGGTTCCTGTACCGAGCAGATGGCCGC 121
QY 661 AACTGTATATATCTAAATATATCCCAACAAAGGCGACGCTTTGTGACCGAACCCCGGAA 720
Db 122 AACTGCATATATCTAAATATATCTCCAGCAAAAGGCGATGCTTTGCTGCAACCCCGGAA 181
QY 721 GAGTATCCAGAGTGCAAGGTTTATGAGAACTGAGGAGCAGCATATGCTGATCCCGTG 780
Db 182 GAGTATCCAGAGTGCAAGGTTTATGAGAACTGAGGAGCAGCATATGCTGATCCCGTG 241
QY 781 AGCATGTCTGAACCTGGAAGAGGTGGATGGGCTGCTCACCTGCTCTCAGTTTAAATTAAC 840
Db 242 AGCAACTCGGAGATGGAAAGGTGGAGCGCTGCTCACCTGCTCTCAGTTTAAATTAAC 301
QY 841 AAGAGGTAGACTCCTGA 858
Db 302 AAGAGGTAGACTCCTGA 319
RESULT 13
AAZ52967
ID AAZ52967 standard; cDNA; 1376 BP.
XX AC AAZ52967;
XX AC AAZ52967;
XX DT 14-MAR-2000 (first entry)
XX

DE Human prostate tumor cDNA library derived EST fragment #110.
XX Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
KW treatment; ds.
XX Homo sapiens.
XX DE19820190-A1.
XX 04-NOV-1999.
XX 28-APR-1998; 98DE-01020190.
XX 28-APR-1998; 98DE-01020190.
XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX P-PSDB; AAY74144, AAY74145, AAY74146.
XX WPI; 1999-621386/54.
XX New human nucleic acid sequences from pancreatic tumors, and related
XX proteins.
XX Claim 2; Page 272; 502pp; German.
XX This invention describes novel polypeptides and their encoding nucleic
XX acids derived from human pancreatic tumor tissue which have cytostatic
XX activity. The sequences are also useful in producing pharmaceutical
XX compositions for treatment of pancreatic tumors. AA252858-253014
XX represent expressed sequence tag (EST) fragments derived from a human
XX pancreatic tumor cDNA library and which encode the proteins represented
XX in AA73814-Y74252
XX Sequence 1376 BP; 294 A; 390 C; 448 G; 244 T; 0 U; 0 Other;
Query Match 25.2%; Score 216.2; DB 2; Length 1376;
Best Local Similarity 54.9%; Pred. No. 2.1e-39;
Matches 450; Conservative 0; Mismatches 363; Indels 6; Gaps 1;
QY 29 TCGCGCGGCGCCAGCGCGTGTGGGGCGCTACCCGAGTGCCTCCAGCAGCGCGC 88
DB 319 TGGCGCGCTCTCCATGCCCTGATCGGGGAGTCCAGAGAGCTGGGTGGGGAG 378
QY 89 TGAGAGCGCCAGGGCGAGAGGTGAGCGTCCCGCGCGGAGCGGACCGACCTCT 148
DB 379 GTGCGGGGCTGGCGCTTCCCGCTCTGATCTGCCCAAAGCTCAAAGGAGCAGGGGTGC 438
QY 149 ACGTGGCGTGTGGCGCAGCAAGCTGGGCTGAGGTGGTGGAGTCCCGCGCGAG 208
DB 439 TGGAGGTAAACTGAGGCAACGACTGGGGCTACAGCTGTAGAACTGCCACCTGAGGT 498
QY 209 GCCTTCCGGAAGTGGTGTGGAGGAGCGTGGCGCGTGTGGAGGAGCGGCGCTCA 268
DB 499 CATTGGCGCTGGACCGCTGTGGCGACACGCGCGTGTATCAAGGGGACCGGCGCTAA 558
QY 269 TCACCGACCGGGGCGCGAGCGGAGGAGGTGACATGATGAAGAAGCATTTAG 328
DB 559 TCACGCGGCGCTGGAGCGGCGCGCTGTAGGCGAGGTGAGTCCGCAAAAGCGCTGC 618
QY 329 AAAAATTCAGCTCAATATAGTAGATGAAGATGAATGAATGAATTTAGATGGCGAG 388
DB 619 AAGACCTGGGCTCCGAAATTTGTGAAATAGGAGACGAGAACGGAGCTGGATGCACTG 678
QY 389 ATGTTTATTCAGGAGAGAAATTTTGTGGGCTTTCCAAAGGACAAATCAACGAG 448
DB 679 ACGTTCTCTTCCACCGGCGGAGTTTTCGTAGGCTCTCCAAATGAGCAATCAACGAG 738
QY 449 GTGCTGAATCTTGGCTGTACTTTTAAAGGACTATGAGTCTCCACAGTCCAGTGGCGAG 508
DB 739 GAGCTGAGATCTGGCGGACACGCTTCGGGACTTCGCCCTCTCCACTGTGCCAGTCTCG 798

QY 509 ATGGGTTGCATTTGAAGAGTTTCTGACGATGCTGGGCTACCTGATCGCAATTGGGT 568
DB 799 GTCCCTCCCACTGCGCGGTCTCTGGCGGACCTGGGGACCTGCGACTGTTGTGGCAGCA 858
QY 569 CTAGTGAATCTGCACAGAGGCGCTTAAGATCATGCAACAGATGAGTGACCCGCTACG 628
DB 859 GCAGCGAGCTGCCCAAAAGGCTGTCGGGCAATGGCAGTGTGACAGATCAACCATATG 918
QY 629 ACAAACTCAGTGCCTGATGACATAGCAGCAAACTGTATATCTAAAT-----ATCC 682
DB 919 CCTCCCTGACCTCCAGATGACGACGCTGTGACTGTCTCTTTCTCGTGGGTGTC 978
QY 683 CCAACAAAGGCGAGTCTTGTGACCGAAACCCCGGAGAGTATCCAGAAAGTCAAAAGG 742
DB 979 CTGGTGTGCCCTTTCTCTGACCGCTGGAGTGGGATCTGCCCAACAGCCAGAGG 1038
QY 743 TTTATGAAACTGAAGACCATATGTGATGCCCGTGAGCATGCTTGAACCTGGAAGG 802
DB 1039 CACTGCAAGACTCTCTGATGTCACTGTTGACCTGTTGCTGCTCTGCTCAGAACTGGAGG 1098
QY 803 TGGATGGCTGCTCACCTGCTGCTCAGTTTAAATTAACA 841
DB 1099 CTGGCGCGGCTCAGCTCCCTCTGCTTGGTGTCTCAGCA 1137
RESULT 14
AAA47656
ID AAA47656 standard; cDNA; 858 BP.
AC AAA47656;
XX
DT 08-NOV-2000 (first entry)
XX
DE Dimethylarginine dimethylaminohydrolase (DDAH2) coding sequence.
XX
KW Dimethylarginine dimethylaminohydrolase; DDAH; DDAH1; DDAH2;
KW arginine deaminase; hyperlipidemia; renal failure; hypertension;
KW restenosis; atherosclerosis; schizophrenia; multiple sclerosis; cancer;
KW ischemia reperfusion injury; septic shock; multi organ failure;
KW arthritis; skin disorders; inflammatory cardiac disease; migraine;
KW infection; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..858
FT /*tag= a
FT /product= "DDAH2"
XX
PN WO200044888-A2.
XX
PD 03-AUG-2000.
XX
PF 26-JAN-2000; 2000WO-GB000226.
XX
PR 26-JAN-1999; 99GB-00001705.
PR 04-JUN-1999; 99GB-00013066.
XX
PA (UNLO) UNIV COLLEGE LONDON.
XX
PI Vallance PJT, Leiper JM, Whitley GSJ, Charles IG;
XX WPI; 2000-543392/49.
DR P-PSDB; AAB01478.
XX
PT Novel methylarginase polypeptides and polynucleotides, used to identify
PT modulators of them, which are used in the treatment of e.g. cancer,
PT hypertension, and bacterial infections.
XX
PS Claim 1; Page 57-58; 68pp; English.
XX
CC Nucleotides encoding methylarginase polypeptides, vectors comprising
CC these nucleotides and the polypeptides themselves can be used in

CC medicaments for the treatment of hyperlipidemia, renal failure,
CC hypertension, stenosis after angioplasty, atherosclerosis,
CC complications of heart failure, schizophrania, multiple sclerosis or
CC cancer. Modulators of the enzyme can be used in medicaments for the
CC treatment of ischemia-reperfusion injury of the brain or heart, cancer,
CC lethal hypertension in severe inflammatory conditions such as septic
CC shock or multi-organ failure, or local and systemic inflammatory
CC disorders including arthritis, skin disorders, inflammatory cardiac
CC disease, migraine, or microbial or bacterial infection
XX
SQ Sequence 858 BP; 150 A; 268 C; 281 G; 159 T; 0 U; 0 Other;

Query Match 25.0%; Score 214.6; DB 3; Length 858;
Best Local Similarity 54.8%; Pred. No. 4.3e-39;
Matches 449; Conservative 0; Mismatches 364; Indels 6; Gaps 1;

QY 29 TCGCCCGGGCCACCCACGCGTGTGGGGCGGTACCCGAGTCGCTGCCACACGCGC 88
DB 23 TGGCGCGCTGCTCCATGCGCTGATCCGGGAGTCCAGAGACCTGGCTCGGGGGAAG 82

QY 89 TGAGAACGGCAAGGGCGAGAGGTGACGCTGCCCGCGGGAACGGCAGCACAGCTCT 148
DB 83 GTGGCGGGGTGGCGCTTCCCGCTCTGATCTGGCAAGCTCAAGGGAGCACGGGTGC 142

QY 149 ACFTGGCGCTGCTGGCGAGCAAGCTGGGGCTGCGAGTGTGGAGCTGCCGGCCGACGAG 208
DB 143 TGGAGGTAACTGAGCAACGACTGGGGCTACAGCTGTAGAACTGCCACCTGAGGAGT 202

QY 209 GCCTTCGCGACTCGCTTCGTCGAGACGCTGGCCGCTGTGTCGAGGAGACGGCCCTCA 268
DB 203 CATTGCCGCTGGGACCGCTGCTGGCACACGCGCGTGATCCAAAGGGGACACGGCCCTAA 262

QY 269 TCACCCGACCGGGGCCCGAGCCGGAGGAGGAGTTGACATGATGAAGAGCATTAG 328
DB 263 TCACGCGGCGCTGAGCGCCCGCTGTAGGCGAGAGTTCGATGAGTCCGCAAGCCCTGC 322

QY 329 AAAAACTTCAGCTCAATATAGTAGAGATGAAGATGAATGAATCAACTTTAGATGGCGAG 388
DB 323 AAGACCTGGGCTCCGAAATGTGGAATAGAGACGAGACGCGCTGATGGCACTG 382

QY 389 ATGTTTATTCACAGGAGAGAAATTTTGTGGGCTTTTCCAAAGAGCAAAATCAACGAG 448
DB 383 ACGTTCCTCTTACCGGCGGGAGTTTTCGTAGGCTCTCCAAATGGACCAATCACCGAG 442

QY 449 GTGCTGAATCTTGGCTGATCTTTAAGGACTATGCTCCACAGTCCAGTCCGAGTCCGAG 508
DB 443 GAGCTGAGATCGTGGCGGACACGTTCCGGGACTTCGCGCTCTCACTGTGCCAGTCTCG 502

QY 509 ATGGTTGCAATTTGAAGAGTTTCTGCAGCATGCTGGGCTTAACCTGATCGCAATTTGGT 568
DB 503 GTCCCTCCACCTGCGGGTCTCTGCGCATGGGGGACCTCGCACTGTTGTGGCAGGCA 562

QY 569 CTAGTGAATCTGCACAGGCGCTTAAGATCATGACAGATGAGTGACCCCGCTACG 628
DB 563 GCACGAGCGTGCACAAAGGCTGTCGGGCAATGGCAGTGTGACAGATCACCCATATG 622

QY 629 ACAAACTCACTGTGCTGATGACATAGACAGCAAACTGTATATATCTAAAT-----ATCC 682
DB 623 CCTCCCTGACCTCCAGATGACGACGCTGCTGCTCTCTTTCTGCTCGCTGGTTCG 682

QY 683 CCAACAAAGGCGAGCTTCTGTCGACCGCAACCCCGGAGAGTATCCGAAAGTGCAGAG 742
DB 683 CTGTTGTGCCCCCTTCTCTCTGACCGCTGGAGTGGGGATCTGCCCAACAGCAGCAGAG 742

QY 743 TTTATGAAAACTGAAGGACCATATGCTGATCCCGGTGAGCATGCTGTAACCTGGAAGG 802
DB 743 CACTGACAGAGCTCTCTGATGTACCTCTGTACCTGTGCTCTGCTCAGACTGAGAGAG 802

QY 803 TGGATGGGCTGCTCACCTGCTGCTCACTTTTAATAACA 841
DB 803 CTGCGCGCGGCTCAGCTCCCTCTGCTTGTGTGCTCAGCA 841

RESULT 15

AAD00677
XX AAD00677 standard; cDNA; 1228 BP.
AC AAD00677;
XX 08-SEP-2000 (first entry)
XX Human Hydrolase protein-2 (HYDRL-2) encoding CDNA.
XX
KW Hydrolase; HYDRL; human; clone 949738; cytosolic; immunosuppressive;
KW antinflammatory; neuroprotective; cerebroprotective; anticonvulsant;
KW nephrotropic; antibody; agonist; antagonist; diagnosis; treatment;
KW prevention; cell proliferative disorder; cancer; inflammation; AIDS;
KW Acquired Immune Deficiency Syndrome; autoimmune/inflammatory disorder;
KW neurological disorder; epilepsy; stroke; medullary sponge kidney;
KW renal disorder; adrenal disorder; adrenoleucodystrophy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 263..1120
FT /*tag= a
FT /product= "Human Hydrolase protein-2"
FT /note= "Derived from PANCNOT05 library"
FT 323..373
FT misc_feature
FT /*tag= b
FT /note= "DNA used in hybridisation and amplification
FT technology for identification of HYDRL sequences"
FT 1067..1117
FT misc_feature
FT /*tag= c
FT /note= "DNA used in hybridisation and amplification
FT technology for identification of HYDRL sequences"
XX
PN WO200028045-A2.
XX
PD 18-MAY-2000.
XX
PF 12-NOV-1999; 99WO-US027009.
XX
PR 12-NOV-1999; 98US-0172256P.
PR 21-MAY-1999; 99US-0135519P.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Tang TY, Hillman JL, Yue H, Lal P, Bandman O, Corley NC;
PI Guegler KJ, Baughn MR, Lu DAM, Azimzai Y, Yang J;
XX
DR WPI; 2000-376557/32.
DR P-PSDB; AAY711104.
XX
PT Novel human hydrolase protein useful for diagnosing, treating and
PT preventing cell proliferative, autoimmune and inflammatory, neurological,
PT renal, adrenal and genetic disorders.
XX
PS Claim 9; Page 94-95; 106pp; English.
XX
CC The present sequence is the human hydrolase protein-2 (HYDRL-2) encoding
CC cDNA, identified in Incyte clone 949738, derived from the PANCNOT05
CC library. It is expressed in reproductive, gastrointestinal, nervous, and
CC cardiovascular tissues. This sequence maps to chromosome 6 within the
CC interval from 42.30 to 45.40 centiMorgans (cM), to chromosome 9 within
CC the interval from 130.40 to 166.50 cM, and to chromosome 16 within the
CC interval from 88.10 to 92.60 cM. HYDRL antibodies, agonists and
CC antagonists are useful for diagnosis, treatment and prevention of
CC disorders associated with altered expression or activity of HYDRL. It
CC includes cell proliferative disorder such as cancer, autoimmune or
CC inflammatory disorders such as inflammation, AIDS, neurological disorder
CC such as epilepsy, stroke, renal disorder such as medullary sponge kidney,
CC adrenal disorder such as adrenoleucodystrophy. The nucleotide sequences
CC are useful for detecting and quantifying gene expression in tissues
XX
SQ Sequence 1228 BP; 237 A; 359 C; 418 G; 214 T; 0 U; 0 Other;

Query Match 25.0%; Score 214.6; DB 3; Length 1228;
Best Local Similarity 54.8%; Pred. No. 4,7e-39;
Matches 449; Conservative 0; Mismatches 364; Indels 6; Gaps 1;

```
QY 29 TCGGCGGCGCCACCCACGCGGTGGTCGGCGCGCTACCCGAGTCGCTCTGCCAGCACGCGC 88
Db 285 TGGGCGCGTGTCTCCCATGCTCCCTGATCCGGGGAGTCCAGAGAGCCTGGCGTCGGGGGAG 344
QY 89 TGAGAAAGCGCAAGGCGGAGGAGTGGACGTCGCCCGCGGAAAGCGGAGACCCAGCTCT 148
Db 345 GTGCGGGGGGTGGCCCTTCGCCGCTCGATCTGGCAAAAGCTCAAAGGAGCACGGGGTGC 404
QY 149 AGTGGGCGGTGTGGGCGAGCAGCTGGGCGTGGAGTGTGGAGCTGCCGCGCGAGAGA 208
Db 405 TGGGAGGTAACTGAGGCAACGACTGGGGCTACAGCTGCTAGAACTGCCACCTGAGGAGT 464
QY 209 GCCTTCGGGACTGCGTCTTCGTGGAGACGTCGGCGGTGGTGTGGAGGAGACGGCCCTCA 268
Db 465 CATTGCGGTGGACCGCTGCTTGGCGCACCGGCGGTGATCCAAAGGAGCACGGCCCTAA 524
QY 269 TCACCCGACCGGGGCGCGAGCCGAGGAGAGGTTGACATGATGAAGAGCAATTAG 328
Db 525 TCACGCGGCGCTGGAGCCCGCTCGTAGGCGAGGTCGATGGAGTCCGCAAGGCCCTGC 584
QY 329 AAAAATTCAGCTCAATATAGTAGATCAAGATGAATGAATGCAACTTTAGATGGCGGAG 388
Db 585 AAGACTGGGGCTCCGAATTGGAATAGGAGACGAGCCGCTGATCCAAAGGAGCACGGCCCTAA 644
QY 389 ATGTTTTATTACAGGCGAGAGATTTTTTTGGGGCTTTCCAAAAGGACAAATCAACGAG 448
Db 645 ACGTTCTCTTACCGGCGGGAGTTTTTCGTAGGCGCTTCCAAATGGACCAATCACCGAG 704
QY 449 GTGCTGAATCTTGGTGTACTTTAAGGACTATGCACTCTCCACAGTCCAGTGGCGAG 508
Db 705 GAGCTGAGATCGTGGCGGACAGTTCGCGGAGTTCGCGCTCTCCACTGTGCCAGTCTCGG 764
QY 509 ATGGGTTCATTTGAAGAGTTTTCGAGCATGGCTGGGCTAACTGATCGCAATTGGGT 568
Db 765 GTCCCTCCACCTGCGCGTCTCTGGGCAATGGGGGACTCGCACTGTTGTGGCAGGCA 824
QY 569 CTAGTGAATCTGCACAGAGGCGCTTAAGATCATGCAACAGATGAGTGACACCGCTAGG 628
Db 825 GCAGCGACGCTGCCAAAAGGCTGTCGGGGCAATGGCAGTGTGACAGATCACCCATATG 884
QY 629 ACAAACTCACTGTGCTGATGACATAGCAGCAAACTGTATATATCTAAAT-----ATCC 682
Db 885 CTTCCCTGACCTCCAGATGACGAGCTGTGACTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 944
QY 683 CCACAAAGGGGACGTCCTTGTGCAACCGAACCCCGGAGAGATCCAGAAAGTGAAGG 742
Db 945 CTGGTGTGCCCCCTTTCCTCTGCAACCGTGGAGGTGGGATCTGCCCAACAGCCAGGAGG 1004
QY 743 TTTATGAGAACTGAGGAGCATATGCTGATCCCGTGAGCATGTCTGAATCTGGAAGG 802
Db 1005 CACTGCAGAGCTCTCTGATGTCACCTGGTACCTGTGTCTCTCTCTCTCTCTCTCTCTCTCT 1064
QY 803 TGGATGGGTGCTCACTGCTGCTCAGTTTAAATTAACA 841
Db 1065 CCGGCGCGGGCTCAGCTCCCTCTGCTTGGTGTCTAGCA 1103
```

Search completed: June 7, 2004, 20:32:46
Job time : 412 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 19:08:46 ; Search time 2569 seconds

(without alignments)
9973.434 Million cell updates/sec

Title: US-09-889-733B-1

Perfect score: 858

Sequence: 1 atggcggcctcgccacc.....acaagaaggtagactcctga 858

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estl:*

9: gb_estl:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_man:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vri:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	717.2	83.6	1109	11	AK010430 Mus muscu
2	644.4	75.1	889	12	BI763990 603049879
3	631.4	73.6	1001	13	BY709864 BY709864
4	582.6	67.9	746	14	CB249709 UI-M-EXO-

5	577.8	67.3	862	13	BQ899468
6	547.8	63.8	760	14	CA512485
7	544.4	63.4	822	9	AI647789 uk43005.x
8	534.2	62.3	656	13	BY742062 BY742062
9	517.8	60.3	522	9	AI751463 cn10b04.x
10	480.8	56.0	698	13	BY754930 BY754930
11	479	55.8	731	10	BB611533 BB611533
12	472.2	55.0	800	9	AA986117 uc81h06.x
13	453.4	52.8	455	29	AY406329 Homo sapi
14	450.2	52.5	455	29	AY406330 Pan trogl
15	445.8	52.0	685	14	CB247154 UI-M-F10-
16	439	51.2	702	13	BY731938 BY731938
17	430.6	50.2	645	29	CG477727 CG477727
18	428	49.9	654	9	AU296050 AU296050
19	425.4	49.6	624	10	AW514143 hd24a02.x
20	413.2	48.2	1035	10	BF579996 602095987
21	411.6	48.0	501	29	CG617513 OSTR11187
22	410.4	47.8	923	12	BI332932 602982027
23	396	46.2	777	14	CD349526 UI-M-FY0-
24	390.4	45.5	860	14	CF548604 AGENCOURT
25	388	45.2	486	9	AI463821 va31a06.y
26	386.4	45.0	806	9	AU051351 AU051351
27	381.8	44.5	681	12	BM963225 UI-M-EQ0-
28	380.4	44.3	569	9	AI265716 uJ09603.x
29	377.2	44.0	776	14	CD751679 AGENCOURT
30	375	43.7	455	29	AY406331 Mus muscu
31	375	43.7	556	12	BM964027 UI-M-EQ0-
32	370.8	43.2	638	13	BY713422 BY713422
33	361.6	42.1	555	9	AI050531 uc83d09.y
34	361.6	42.1	817	12	BI221174 602939511
35	359	41.8	792	14	CD750695 AGENCOURT
36	357.6	41.7	547	9	AI661738 va31a06.x
37	357.6	41.7	897	14	CK019367 AGENCOURT
38	357	41.6	432	10	BB84630 UI-M-BH1-
39	356	41.5	550	9	AA986603 uc81h06.y
40	352.8	41.1	356	12	EG977444 RCS-CI016
41	344.6	40.2	571	14	CB522710 UI-M-GK0-
42	344.6	40.2	571	14	CB522972 UI-M-GK0-
43	338.6	39.5	539	9	AW012296 um07b05.y
44	338.6	39.5	720	9	AV167374 AV167374
45	338.2	39.4	802	14	CA588435 hab63f06.

ALIGNMENTS

RESULT 1
AK010430
LOCUS 1109 bp mRNA linear HTC 20-SEP-2003
DEFINITION Mus musculus ES cells cDNA, Riken full-length enriched library, clone:2410006N07 product:dimethylarginine dimethylaminohydrolase 1, full insert sequence.
ACCESSION AK010430
VERSION AK010430.1 GI:12845867
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
Carninci, P. and Hayashizaki, Y.
AUTHORS High-efficiency full-length cDNA cloning
TITLE Meth. Enzymol. 303, 19-44 (1999)
JOURNAL 99279253
MEDLINE 10349636
PUBMED 10349636
REFERENCE 2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
AUTHORS Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
TITLE Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL 20493374
MEDLINE 11042159
PUBMED 11042159

QY 841 AAGAAGGTAGACTCTCTGA 858
|||||
Db 977 AAGAAGATAGACTCTCTGA 994

RESULT 2
BI763990 889 bp mRNA linear EST 25-SEP-2001
LOCUS 603049879F1 NTH_MGC_116 Homo sapiens cDNA clone IMAGE:5189970 5',
DEFINITION mRNA sequence.
ACCESSION BI763990
VERSION BI763990.1 GI:15755568
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 889)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11474 row: n column: 19
High quality sequence start: 3
High quality sequence stop: 847.
Location/Qualifiers
1..889
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5189970"
/lab_host="DH10B"
/clone_lib="NIH_MGC_116"
/note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPOK6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (in vitro). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."

FEATURES

source
1..889
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5189970"
/lab_host="DH10B"
/clone_lib="NIH_MGC_116"
/note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPOK6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (in vitro). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 75.1%; Score 644.4; DB 12; Length 889;
Best Local Similarity 92.9%; Pred. No. 4.7e-110;
Matches 732; Conservative 0; Mismatches 46; Indels 10; Gaps 5;
QY 1 ATGGCGGCGCTCGGCCACCCCTTCGCGTTCGCGCGGCCACCCACCGCGTGTGGGCGG 60
Db 102 ATGGCGGCGCTCGGCCACCCCTTCGCGTTCGCGCGGCCACCCACCGCGTGTGGGCGG 161
QY 61 CTACCCGAGTGGTCTGCCAGACCGGCTGAGAGGCCCAAGGGCGAGAGGTGGACGTC 120
Db 162 CTACCCGAGTGGTCTGCCAGACCGGCTGAGAGGCCCAAGGGCGAGAGGTGGACGTC 221
QY 121 GCCCGCGCGGAACCGGACGACCGACCGCTCTACGTGGGCGTCTCGGCGAGCAAGCTGGGGCTG 180
Db 222 GCCCGCGCGGAACCGGACGACCGACCGCTCTACGTGGGCGTCTCGGCGAGCAAGCTGGGGCTG 281
QY 181 CAGGTGGTGGAGTGC CGGCGGACGAGAGCTTCGCGACTGCGTCTTCGTGGAGACGTG 240
Db 282 CAGGTGGTGGAGTGC CGGCGGACGAGAGCTTCGCGACTGCGTCTTCGTGGAGACGTG 341

QY 241 GCCGTGTGTGGAGGAGACGGCCCTCATCCCGACCCCGGGCGCGAGCGGAGAG 300
Db 342 GCCGTGTGTGGAGGAGACGGCCCTCATCCCGACCCCGGGCGCGAGCGGAGAG 401
QY 301 GAGGTTGACATGATGAAGAAGCAATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
Db 402 GAGGTTGACATGATGAAGAAGCAATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAA 461
QY 361 GATGAATAATCAACTTTAGATGCGGAGATGTTTTATTTCAGGCGAGAGATTTTTTGTG 420
Db 462 GATGAATAATCAACTTTAGATGCGGAGATGTTTTATTTCAGGCGAGAGATTTTTTGTG 521
QY 421 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTTGGCTGATACCTTTAAGGAC 480
Db 522 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTTGGCTGATACCTTTAAGGAC 581
QY 481 TATGCAAGTCTCCACAGTGCAGTGGCAGATGGGTTCGATTTTGAAGAG- TTTCTGACGAT 539
Db 582 TATGCAAGTCTCCACAGTGCAGTGGCAGATGGGTTCGATTTTGAAGAGTTTCTGACGAT 641
QY 540 GGTGGGCGCTAACTGATCGCAATTTGGTCTAGTGAATCTGACAGAGGCGCTTAAGAT 599
Db 642 GGTGGGCGCTAACTGATCGCAATTTGGTCTAGTGAATCTGACAGAGGCGCTTAAGAT 701
QY 600 CATGCAACAGA-TGAGTGACACCGCTACGAC-AAACTCACTGT-GCCTGATGACATAGC 656
Db 702 CATGCAACAGATTGAGTGACCGCTACGACAAACTCACTGTGGCTGATGACCTTAG 761
QY 657 AGCAACTGTATATCT-----AAATATCCCAACAAAGGCGAGTCTTGTGACCG 710
Db 762 CAGCAAACTGTATATCTTAAATTTATCCCAACAAAGGCGCGGTCTTGTGACG 821
QY 711 AACCCCGGAAGAGTATCCAGAAAGTGCAAAAGTTTATGAGAACTGAAGACCATATGCT 770
Db 822 GRACCCGAGAGATTCCAGAAAGTGCAAAAGTTTATGACAACTTGAGGACCTTAAGT 881
QY 771 GATCCCGG 778
Db 882 GATCCCGG 889

RESULT 3
BI709864 1001 bp mRNA linear EST 16-DEC-2002
LOCUS BY709864 RIKEN full-length enriched, ES cells Mus musculus CDNA
DEFINITION Clone 2410006N07 5', mRNA sequence.
ACCESSION BY709864
VERSION BY709864.1 GI:27121074
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1001)
AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojbori, T., Baldarelli, R., Hill, D. P., Bult, C.,
Hume, D. A., Quackenbush, J., Schriml, I. M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V.,
Ciothica, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A.,
Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,
Gariboldi, M., Glasi, C., Godzik, A., Gough, J., Grimmond, S.,
Guatini, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Redziarski, R. M., King, B. L., Konagaya, A.,
Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ribg, B. Z., Ringwald, M.,
Sandelin, A., Schneider, C., Sempie, C. A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,

Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilm, B.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalon, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakasume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Akazawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingagawa, A., Tasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, K.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

223354683
12466851

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-32 Suihiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-resesgsc.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,
Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,
Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,
Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,
Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,
Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,
Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
SOURCE

Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57Bl/6J"
 /db_xref="taxon:10090"
 /clone="241006N07"
 /cell_type="ES cells"
 /lab_host="SOLR"
 /clone_libs="RIKEN full-11
 /notes="Site 1: xhoi; Site
 prepared and sequenced in
 Project of Genome Explor
 Genomic Sciences Center
 RIKEN. Division of Exper
 contributed to prepare m
 primed with a primer [5'
 GAGACGAGAGAAGGATCCAAAGGAT
 prepared by using trehal
 transcriptase and subsequ
 cap-trapper. cDNA went
 to Rot = 5.0 and subtra
 cDNA was prepared with t
 GAGACGAGATTCGAGTAATTC

ORIGIN

Query Match 73.6%; Score 631.4; DB 13; Length 1001;
Best Local Similarity 89.3%; Pred. No. 1.3e-107;
Matches 702; Conservative 0; Mismatches 82; Indels 2; Gaps 2;

QY 1 ATGGCGGCGCTCGGCCACCCCTCCGCCTTCGCGCGGGCCACCACCGCGTGGTGCGGGCG 60
DB |||||
QY 61 CTATCCCGAGTCGCTCTGCGCAGCACGCGCTGAGAAGCGCAAGGCGAGGAGTGGACGTG 120
DB |||||
QY 194 CGCGCCGAGTCCTGTGTCGGCCACGCGCTGAGCGCTCGCAGGGCGAGAGTGGATTTC 253
DB |||||
QY 121 GCCCGCGGAACGCGACGACC-AGCTCTACG-TGGGCGTGTGGCAGCACAAGCTCGGGC 178
DB |||||
QY 254 GTTCGCGCGAGCGCCAGCACCGAGCTCTACGCTGGGCGTGTGGGCGACAAGCTGGGGC 313
DB |||||
QY 179 TGCAGTGTGTGAGTGTCCCGGCGCGACGAGCGCTTCGCGACTGCGTCTTCGTGGAGGACG 238
DB |||||
QY 314 TGCAGGTGTGCAGCTGCCCGCGGACGAGCGCTGCCGACTGCGTGTTCGTGGAGGACG 373
DB |||||
QY 239 TGCCGTGTGTGCGAGGAGAGCGCCCTCATACCCGACCCGGGGCGCGAGCCGCGAGGA 298
DB |||||
QY 374 TGCCGTGTGTGCGAGGAGAGCGCCCTCATACCCGCCCCGGGGCGCCGACGCGCAGGA 433
DB |||||
QY 299 AGGAGGTTCACTGATGAAGAAGACATTAGAAAACCTTCAGCTCAATATAGTAGAGATGA 358
DB |||||
QY 434 AGAGGTTGACATGATGAAGAAGGCTTTCGAAAAAAGCTTCAGCTCAACATAGTAGAGATGA 493
DB |||||
QY 359 AAGATGAAATGCAACTTTAGATGCGCGAGATGTTTTATTTCAGGCGAGAGAAATTTTTCG 418
DB |||||
QY 494 AAGATGAAATGCAACTTTGGATGTTGGGAGCGTCTCTATTTCACAGGCGAGAGAAATTTTTCG 553
DB |||||
QY 419 TGGGCGTTTCCAAAGGACAAATCACAGAGGTGCTGAAATCTTGGCTGATCTTTTAAGG 478
DB |||||
QY 554 TGGGCGTTTCCAAAGAACAAATCAAAGAGTGTGAAATCTTGGCTGATCTTTTAAGG 613
DB |||||
QY 479 ACTATGAGTCTCTCCAGTGCAGTGGCAGATGGGTTCATTTGAAGATTTCTGCAGCA 538
DB |||||
QY 614 ACTAGCAGTCTCTACAGTCCCTGTGGCGGATCTCTTGCATTTAAAGAGTTTCTGCAGCA 673
DB |||||
QY 539 TGGCTGGGCTAACCTGATCGCAATTTGGCTAGTGAATCTGCACAGAGGCCCTTAAGA 598
DB |||||
QY 674 TGGCGGACCCAACTGTGTAATAGGGTCCAGCGAATCTGCACAGAGGCGCTCAAGA 733
DB |||||
QY 599 TCATGCAACAGATGAGTGACCCGCTAGCAAACTCACTGTGCTGTATGACATAGCAG 658
DB |||||
QY 734 TCATGCAACAGATGAGTGACCATCGTTATGACAAGCTCACTGTACCGACGACATGCGCG 793
DB |||||
QY 659 CAACTGTATATATCTAATAATCCCAACAAGGGCAGCTTCTGTGCACCGAACCCCG 718
DB |||||
QY 794 CCAACTGCATATATCTAAATATATCCCAGCATAGGGCATGTCTTGTGCTGCACTGAACCCAG 853
DB |||||
QY 719 AAGAGTATCCAGAAAGTGCAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCGG 778
DB |||||
QY 854 AAGAGTACCCAGAAAGGCAACGCTATGAGNNACTCAAGGACCATCTACTGATCTCTG 913
DB |||||

QY 779 TGAGCA 784
DB |||||
QY 914 TGAGCA 919
DB |||||

RESULT 4
CB249709
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CB249709
UT-M-EXO-by1-1-07-0-UI.r1 NIH BMAP_EXO Mus musculus cDNA clone
IMAGE:5719830 5', mRNA sequence.
CB249709
CB249709.1 GI:28389411
EST.
Mus musculus (house mouse)
Mus musculus

746 bp mRNA linear EST 15-JUL-2003

RESULT 4
CB249709
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CB249709 746 bp mRNA linear EST 15-JUL-2003
 UI-M-EX0-by1-1-07-0-UI.r1 NIH_BMAP_EX0 Mus musculus cDNA clone
 IMAGE:5719830 5', mRNA sequence.

ACCESSION	CB249709
VERSION	CB249709.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 746)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.

FEATURES
 source

Location/Qualifiers
 1. .746
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5719830"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP EX0"
 /note="Organ: Brain; Vector: pYX-Asc; Site: 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTCCGTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 67.9%; Score 582.6; DB 14; Length 746;
 Best Local Similarity 89.8%; Pred. No. 1.5e-98;
 Matches 624; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
 QY 164 GCAGCAAGCTGGGGCTGCAGGTGGTGGAGCTCCCGCGACGAGAGCCTTCGGACTCG 223
 DB 1 GCAGCAAGCTGGGGCTGCAGGTGGTGGAGCTCCCGCGACGAGAGCCTTCGGACTCG 60
 QY 224 TCTTCGTGAGGACGTGGCGCTGGTGTGCGAGGACGCGCCCTCATCCCGACCCCGGG 283
 DB 61 TGTTCGTGAGGACGTGGCGCTGGTGTGCGAGGACGCGCCCTCATCCCGCGGG 120
 QY 284 CCGCGAGCGGAGGAGGAGGTTGATGATGAAGAAGCATAGAAAACCTTCAGTCTCA 343
 DB 121 CCGCGAGCGGAGGAGGAGGTTGATGATGAAGAAGCATAGAAAACCTTCAGTCTCA 180
 QY 344 ATATAGTAGAGTGAAGATGAATGAACATTTAGATGGCGGAGATGTTTATTTCAG 403
 DB 181 ACATAGTAGAGTGAAGATGAATGAACATTTAGATGGCGGAGATGTTTATTTCAG 240
 QY 404 GCAGAGATTTTGTGGCCCTTCCAAAGACAAATCAACAGAGTGCTGAATCTTGG 463
 DB 241 GCAGAGATTTTGTGGCCCTTCCAAAGACAAATCAACAGAGTGCTGAATCTTGG 300
 QY 464 CTGATACCTTTTAAGACTATGAGTCTCCACAGTGCAGTGGCAGATGGTTGATTTGA 523
 DB 301 CTGATACCTTTTAAGACTATGAGTCTCCACAGTGCAGTGGCAGATGGTTGATTTGA 360

QY 524 AGAGTTTCTCAGCATGGCTGGCGCTTAACCTGATCGAATTGGTCTAGTGAATCTGCAC 583
 DB 361 AGAGTTTCTCAGCATGGCTGGCGCGACCCCAACCTGATTGCAATAGGCTCAGCGAATCTGCAC 420
 QY 584 AGAAGGCGCTTAAGATCATGCAACAGATGAGTGACCAACCGCTACGACAAACTCAGTGTGC 643
 DB 421 AGAAGGCGCTCAAGATCATGCAACAGATGAGTGACCACTCGTTATGACAAGCTCACTGTAC 480
 QY 644 CTGATGATCATGAGCAAACTGTATATATCTAAATATCCCAACAAAGGCGAGCTTGTGC 703
 DB 481 CCGAGCATGTCGGCGCAACTGATATATCTAAATATCCCAACAAAGGCGAGCTTGTGTGC 540
 QY 704 TGCAACCAACCCCGAGAGTATCCAGAAAGTGCAGAAAGTTTATGAGAAACTGAAGGACC 763
 DB 541 TGCACCGAACCAGAGAGTACCCAGAAAGCGCAGAGGTCTATGAGAACTCAAGGACC 600
 QY 764 ATATGCTGATCCCGTGGAGCATGTCTGAACTGAAAGGTGGATGGGCTGCTACCTGCT 823
 DB 601 ATCTACTGATCCCTGTGAGCAACTCGGAGATGAAAGGTGGAGCGGCTTGTCACTGCT 660
 QY 824 GCTCAGTTTAAATTAACAAGAAGTGAGACTCTCTGA 858
 DB 661 GCTCCGTTTATTAACAAGAATAGACTCTCTGA 695
 RESULT 5
 BQ899468 862 bp mRNA linear EST 16-AUG-2002
 LOCUS AGENCOURT_8750357 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6332815
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ899468
 VERSION BQ899468.1 GI:22291482
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 862)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.
 cDNA Library Preparation: ResGen, Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: L14M13791 row: a column: 08
 High quality sequence stop: 704.
 FEATURES
 source
 1. .862
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6332815"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 130"
 /note="Organ: oocytes; Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."
 ORIGIN
 Query Match 67.3%; Score 577.8; DB 13; Length 862;
 Best Local Similarity 89.6%; Pred. No. 1.3e-97;
 Matches 621; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
 QY 1 ATGCGCGGCTCGGCACACCCCTCGGCTTCGGCGGCGCACCCACCGCTGGTGGCGGG 60

Db 131 ATGGCGGCTCTGCGCCACCCCTCGCTTCTGGCGGCGCCACCCAGCGCTGTGGCGGCT 190
Qy 61 CTACCCGAGTCTGCTGCGACAGCGCTGAGAGCCCAAGGCGGAGGAGTGGACGTC 120
Db 191 CCGCCGAGTCTGCTGCGCGCGCTGAGGCGCTGCGAGGCGGAGGAGTGGATTTC 250
Qy 121 GCCCGCGGAGCGGACAGCAGCTCTACGTGGGCGTGTGGGCGGCTGTGGGCGGAGCTGGGCGTG 180
Db 251 GCTGGCGGAGCGGACAGCAGCTCTACGTGGGCGTGTGGGCGGAGCTGGGCGTG 310
Qy 181 CAGTGTGGAGCTGCGCGGCGGACAGAGCTTCCGAGCTGCGTTCCTGTTGGAGGAGCTG 240
Db 311 CAGTGTGGAGCTGCGCGGCGGACAGAGCTTCCGAGCTGCGTTCCTGTTGGAGGAGCTG 370
Qy 241 GCGTGTGTGGAGGAGCGGCTCATCAGGAGCGGCGGCGGCGGCGGCGGCGGAG 300
Db 371 GCGTGTGTGGAGGAGCGGCTCATCAGGAGCGGCGGCGGCGGCGGCGGCGGAG 430
Qy 301 GAGTGTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
Db 431 GAGTGTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 490
Qy 361 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
Db 491 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 550
Qy 421 GCGCTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
Db 551 GCGCTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 610
Qy 481 TATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
Db 611 TACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 670
Qy 541 GCTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Db 671 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 730
Qy 601 ATGCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
Db 731 ATGCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 790
Qy 661 AACTGTATATCTTAAATATCCCAACAAAGG 693
Db 791 CACTGCATATATCTTAAATATCCCAACAAAGG 823

RESULT 6
CAS12485
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

CAS12485
UI-R-FU0-cpz-1-06-0-UI.r1 UI-R-FU0 Rattus norvegicus cDNA clone
EST
CAS12485.1 GI:25003439
Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 760)
Bonaldi,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250

Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.reagen.com).
Seq primer: M13 REVERSE.
Location/Qualifiers
1..760
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-FU0-cpz-1-06-0-UI"
/lab_stages="embryo"
/tissue_type="embryo"
/clone_lib="UI-R-FU0"
/note="Vector: pX-Asc; Site 1: Ecor I; Site 2: Not I;
UI-R-FU0 is a cDNA library containing the following
tissue(s): rat embryo. The library was constructed
according to Bonaldi, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an Ecor I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dr)18 tail. The
sequence tag for this library is CATCTACT. This library
was created for the University of Iowa Program for Rat
Gene Discovery and Mapping (Val Sheffield, Bento Soares
and Tom Casavant)"

ORIGIN

Query Match 63.8%; Score 547.8; DB 14; Length 760;
Best Local Similarity 89.8%; Pred. No. 4.8e-92;
Matches 588; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
Qy 204 CGAGAGCCTTCGGAGTGGCTTCTGTTGAGAGACGTCGCGCGTGTGGAGAGACGCG 263
Db 1 CGAGAGCCTTCGGAGTGGCTTCTGTTGAGAGACGTCGCGCGTGTGGAGAGACGCG 60
Qy 264 CTTATCATCCCGACCG 323
Db 61 CTTATCATCCCG 120
Qy 324 ATTAGAAACCTTCAGCTCAATATATAGTACAGATGAAGATGAATGCAACTTTAGATGG 383
Db 121 TTGGAAACCTTCAGCTCAATATAGTACAGATGAAGATGAATGCAACTTTAGATGG 180
Qy 384 CGAGAGTGTTTTATTACAGGCGAGAGATTTTGTGGCGCTTCCAAAAGGACAAATCA 443
Db 181 TGGGAGCGTCTATTACAGGCGAGAGATTTTGTGGCGCTTCCAAAAGGACAAATCA 240
Qy 444 ACAGGTGTGAAATCTGGCTGATCTTTTAAAGCACTTGCAGTCTGCAGTGCAGTGCAGT 503
Db 241 ACAGGTGTGAGATCTTGGCTGATCTTTTAAAGCACTTGCAGTCTGCAGTGCAGTGCAGT 300
Qy 504 GGCAGATGGTGTGCATTTGAAGAGTTTTCGACAGTGTGGCGCTTAACCTCATCGCAAT 563
Db 301 GGCAGATTTTGTGCATTTTAAAGAGTTTTCGACAGTGTGGCGCTTAACCTCATCGCAAT 360
Qy 564 TGGGTCTAGTGAATCTGCACAGAGCGCTTAAAGATCATGCNAAGATGAGTGACACCG 623
Db 361 AGGTCCAGTGAATCTGCAGAGCGCTTAAAGATCATGCNAAGATGAGTGACACCG 420
Qy 624 CTACGACAACTCACTGTGGCTGTGATGATAGCAGCAAACTGTATATCTATAATATCCC 683
Db 421 TTATGACAGCTCACTGTACCGGACATGATGCCCGCACTGTATATTTAATATCCC 480

Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gusticich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Matsuda, L., Marchionni, L.S., McKenzie, L., Miki, H., Nagashima, T., Nunata, K., Okido, T., Pavan, W.J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wallestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itch, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E., and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp

URL: http://genome.gsc.riken.go.jp/

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohashi, N., Saito, R., Sakazume, N., Sato, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takada, Y., Waki, K., Watanabe, K., Muramatsu, M., and Hayashizaki, Y.

Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

RNA was provided by Dr. Akira Nakagawa (Div. of Biochemistry, Chiba Cancer Center Research Institute, 666-2 Nitona, Chuo-ku, Chiba, 260-8717 Japan) whose assistance is gratefully acknowledged. Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

Location/Qualifiers

1. 656

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="7120491N07"

/sex="mixed"

/tissue_type="sympathetic ganglion"

FEATURES

source

/dev_stages="2 days neonate"

/lab_host="DH10B"

/clone_lib="RIKEN full-length enriched, 2 days neonate sympathetic ganglion"

/note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCCAAGAGCTCTTTTTTTTTTNN 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCTCGAGTAAATTAATATCCCTCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pluescript KS(+) after bulk excision from Lambda PUC 1. Cloning sites, 5' end: Sali; 3' end: BamHI. Host: DH10B. RNA was provided by Akira Nakagawa, Div. of Biochemistry, Chiba Cancer Center Research Institute, 666-2 Nitona, Chuo-ku, Chiba, 260-8717 Japan, whose assistance we gratefully acknowledge."

ORIGIN

Query Match 62.3%; Score 534.2; DB 13; Length 656;

Best Local Similarity 89.5%; Pred. No. 1.6e-89;

Matches 586; Conservative 0; Mismatches 68; Indels 1; Gaps 1;

182 AGTGTGTGAGCTGCGGCGGACGAGAGCTTCCGAGCTGCTTCGTGGAGGACGTGG 241

1 AGTGTGTGAGCTGCGGCGGACGAGAGCTTCCGAGCTGCTTCGTGGAGGACGTGG 60

242 CGTGTGTGCGAGAGAGCGGCTCATCCCGCGCGGCGGCGGCGGCGGAGG 301

61 CGTGTGTGCGAGAGAGCGGCTCATCCCGCGCGGCGGCGGCGGCGGAGG 120

302 AGTGTGACATGATGAAAGAGCATTAGAAAACTTTCAGCTCAATATAGTAGAGTGAAG 361

121 AGTGTGACATGATGAAAGAGGCTTTGGAAAACTTTCAGCTCAATATAGTAGAGTGAAG 180

362 ATGAAATCAACTTTAGATGCGGAGATGTTTTTATTCACAGGCGAGAGATTTTGTGG 421

181 ATGAAATCAACTTTGAGTGTGGGAGCTCTTATTCACAGGCGAGAGATTTTGTGG 240

422 GCCTTTCCAAAAGGACAAATCAACGAGGTCGTGAATCTTGGCTGATATCTTTAAGGACT 481

241 GCCTTTCCAAAAGGACAAATCAACGAGGTCGTGAATCTTGGCTGATATCTTTAAGGACT 300

482 ATGCACTCTCCACAGTGCAGTGGCAGATGGGTGCTTGAAGAGTTTCTCAGCATGG 541

301 AGCAGTCTCTACAGTCCCTGTGGCGGATTTTGCATTTAAAGAGTTTCTCAGCATGG 360

542 CTGGGCTTAACCTGATCGCAATTGGTCTAGTGAATCTGCACAGAGGCGCTTAAGATCA 601

361 CGGACCCAACTGATGCAATAGGTCAGCGATCTGCACAGAGGCGCTTAAGATCA 420

602 TGCAACAGATGAGTGACCAACCGCTACGAAACTCACTGTGCTGATGATGATGAGCA 661

421 TGCAACAGATGAGTGACCAACCGCTACGAAACTCACTGTGCTGATGATGAGCA 480

662 ACTGTATATCTTAATATCCCAACAGGCGAGTCTTGTGACCGAAACCCCGGAAG 721

481 ACTGTATATCTTAATATCCCAACAGGCGAGTCTTGTGACCGAAACCCCGGAAG 539

722 AGTATCCAGAAAGTGCAAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCGTGA 781

540 AGTATCCAGAAAGTGCAAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCGTGA 599

782 GCATGCTGAAGTGCAAAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCGTGA 836

600 GCAACTCGAGATGGAAGAGGTCGCTTGTCTCCTGCTGCTCCGCTTTATT 654

RESULT 9	AI751463	522 bp	mRNA	linear	EST 20-JUN-2002
LOCUS	AI751463				
DEFINITION	cln0b04.x1 Normal Human Trabecular Bone Cells Homo sapiens cDNA				
ACCESSION	AI751463				
VERSION	AI751463.1	GI:5129727			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 522) Jia, L.B., Young, M.F., Touchman, J.W., Bouffard, G.G., Beckstrom-Sternberg, S.M., Green, E.D., Powell, J.I., Yang, L.M., Robey, P.G., Hortsch, R.N. and Francomano, C.A.				
TITLE	SCAP: The Skeletal Genome Anatomy Project				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Libin Jia Medical Genetics Branch National Human Genome Research Institute 10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA Tel: 301-402-4877 Fax: 301-496-7157 Email: libin@helix.nih.gov DNA Sequencing and Analyses by National Institutes of Health Intramural Sequencing Center (NISC). Plate: 10 row: b column: 04 Seq primer: -21M13 forward primer (ABI). Location/Qualifiers 1. .522 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="NHTBC.cln0b04" /sex="Female" /tissue_type="Bone" /cell_type="Trabecular Bone Cells" /lab_host="SURE" /clone_lib="Normal Human Trabecular Bone Cells" /note="Organ: Hip; Vector: pBluescript; Site1: EcoRI; Library constructed by Dr. Marian Young and Dr. Pamela Gehron Robey (NIDCR)"				
FEATURES					
ORIGIN					
Query Match	60.3%;	Score 517.8;	DB 9;	Length 522;	
Best Local Similarity	99.6%;	Pred. No. 1.7e-86;			
Matches 519;	Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	322	GCATTGAGAAAACTTCAGCTCAATATATAGTACAGATGAAGATGAATGCAACTTTAGAT	361		
Db	2	GCACGAGAAAACTTCAGCTCAATATATAGTACAGATGAAGATGAATGCAACTTTAGAT	61		
QY	382	GGCGGAGATCTTTTATTACAGGCAGAGATTTTGTGGGCGCTTTCCTCAAAAGGACAAAT	441		
Db	62	GGCGGAGATGTTTATTACAGGCAGAGATTTTGTGGGCGCTTTCCTCAAAAGGACAAAT	121		
QY	442	CACAGAGGTCTGAAATCTTGGCTGTATCTTTTAAGGACTATGAGTCTCCACAGTGCCA	501		
Db	122	CACAGAGGTCTGAAATCTTGGCTGTATCTTTTAAGGACTATGAGTCTCCACAGTGCCA	181		
QY	502	GTGCGAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCTTAACCTGATGCCA	561		
Db	182	GTGCGAGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGGCTGGGCTTAACCTGATGCCA	241		
QY	562	ATTGGGTTAGTGAATCTGCACAGAGGCCCTTTAAGATCATGCAACAGATGAGTACCAC	621		
Db	242	ATTGGGTTAGTGAATCTGCACAGAGGCCCTTTAAGATCATGCAACAGATGAGTACCAC	301		
QY	622	CGCTACGACAACTCACTGTGGCTGTATGACATGACAGCAAACTGTATATCTAAATATC	681		

/note="Site 1: XhoI; Site 2: Sati; cdna library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cdna was primed with a primer [5']
GAGAGAGAAGATCCACAGACTCTTTTCTTTTCTTTTCTT
prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cdna went through one round of normalization. Rot = 10.0 and subtraction to Rot = 50.0. Second strand cdna was prepared with the primer adapter of sequence [5']
GAGAGAGAAGATTCGAGTTAAATTAATTCCTCCCTCCCCC 3']. cdna was cleaved with XhoI and Sati."

ORIGIN

Query Match	55.8%	Score 479	DB 10	Length 731
Best Local Similarity	88.8%	Pred. No. 3.3e-79		
Matches 529	Conservative	0	Mismatches 66	Indels 1
			Gaps 1	

1	ATGGCGGCTCGGCCACCCCTCGCGTTCGGCGGGCCACCCACGCCGTGTGCGGGCG	60
134	ATGGCGGGCTCGGCCACCCCTCGCGTTCGGCGGGCCACCCACGCCGTGTGCGGGCT	193
61	CTACCGAGTCGCTTCGCCAGCAGCGCTCAGAAAGCCCAAGGCGAGAGGTGGACGTC	120
194	CCGCCGAGTCCCTGTGTCCGCCACGGCTCAGAGCGCTCGCAGGCGGAGAGGTGGATTC	253
121	GCCCGCGCGAAACGGCAGCACACAGCTCTACGTGGGGCGTGTGGCACAAGCTGGGGCTG	180
254	GCTCGCGCCAGCGCCACGACGAGCTCTACGTGGGGCGTGTGGCACAAGCTGGGGCTG	313
181	CAGGTGGTGAGCTGCCGGCGACGAGAGCCTTCGGACTCGCTTCGTGTGGAGGACGTG	240
314	CAGGTGGTCAGCTGCCGGCGACGAGAGCTGCCGACTCGCTGTTCGTGTGAGGACGTG	373
241	GCCTGGTGTGCGAGGAGACGGCCCTCATCACCGACCCGGGGCGCCGACCCGAGGAAG	300
374	GCGTGTGTGCGAGGAGACGGCCCTCATCACCGCCCGGGCGGCCACGCCGAGGAAG	433
301	GAGGTTGACATGATGAAGAAGCATTTAGAAAACTTCAGTCAATATAGTAGAGATGAAA	360
434	GAGGTTGACATGATGAAGAAGGCTTTGGAAAACTTCAGTCAACATAGTAGAGATGAAA	493
361	GATGAAAATGCAACTTTAGATGCGGAGATGTTTTATTCACAGGCAGAGAAATTTTGTG	420
494	GATGAAAATGCAACTTTGGATGCTGGGGACGTCCTATTTCAGGCAGAGAAATTTTGTG	553
421	GGCCTTTCCAAAGGACAAATCAACGAGTGTGAAATCTTTGGCTGATCTTTTAAGGAC	480
554	GGCCTTTCCAAAGAACAAATCAACGAGTGTGAAATCTTTGGCTGATCTTTTAAGGAC	613
481	TATGCAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCATTTGAAGAGTTTCTGCAG-CAT	539
614	TACGCAGTCTTACAGATCCTGTGGCCGATTCITTTGCAITTTAAGAGATTCGACCCAT	673
540	GGTGGGCTTAACCTGATCGCAATGGGCTTAGTGAACTGTGCACAGAGGCCCTTA	595
674	GGCCGACCCAACTGATTGCATTAAGAGTTCACGGAATCTGCACAGATGCCCTTCA	729

RESULT 12	
AA986117/c	
LOCUS	800 bp mRNA linear EST 28-MAY-1998
DEFINITION	ucsh106.xl sugano mouse kidney mkoa Mus musculus cDNA clone
	IMAGE:1432091.3, similar to TR:O08557 O08557
	N-G-N-G-DIMETHYLARGININE DIMETHYLAminoHYDROLASE. ;, mRNA sequence.
ACCESSION	AA986117
VERSION	AA986117.1 Gi:3167992
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus

Eukaryota: Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1. Murus 1. Biller, L. Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisell, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, B., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennan, G., Soares, B., Wilson, R. and Waters, R.

TITLE	JOURNAL	COMMENT

WASHINGTON UNIVERSITY SCHOOL OF MEDICINE
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 FAX: 514 286 1810

MGI:916159
 Possible reversed clone: similarity on wrong strand
 Seq primer: primer name ambiguous
 High quality sequence stop: 488.

FEATURES source

1. /organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1432091"
/sex="female"
/dev_stages="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse"
/note="Organ: Kidney; Vector:
(CACTGTGTC); site 2: Drago
was primed with an oligo
[ATGGCGCCCTTTTCTTTTTTTT]
ligated to a DraIII adap-
ter and cloned into distinct
vector (5' site CACTGTGTC)
be used to isolate the cDNA
performed to exclude frag-
ments constructed by Dr. Sumio
Institute of Medical Sci-
sequencing: 5' end primer
primer CGAATCGAGTGCAGCT

ORIGIN

Query Match	55.0%;	Score 472.2;	DB 9;	Length 800;
Best Local Similarity	86.4%;	Pred. No. 6,3e-78;		
Matches 579;	Conservative 0;	Mismatches 84;	Indels 7;	Gaps 5;
196 QY	CCGCCGCCGACGAGCGCTTCCGGACTG-	-CGTCTCGTGGAGACAGTGGCGCGTGGTGTGCG	253	
800 Db	CCGCCGCCGACGAGCGCTGCCGACTGGCGTGT	TTTGTGAAGAACGTGCCCGTCGTGTGCG	741	
254 QY	AGGAGA-CGGCCCTCATCACCGACCCGGGG-	-CGCCGAGCCGGAGGAAGGA-GGTTGAC	309	
740 Db	AGGAGACGGCGCTTCATACCCCGCCCGGGCCGCCAC	CGCCGCAAGAAAGGAGGGTTGCC	681	
310 QY	ATGATGAAGAAGACATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAGA	TGAAGATGAAAAAT	369	
680 Db	ATGATGAAGAGGCTTTGAAAAAACTTCAGCTCAACATAGTAGAGATGAAGA	TGAAAAAT	621	
370 QY	GCAACTTTAGATGGCGGAGTCTTTATTCACAGGCAGAGATTTTT	-TGTCGGCCCTTTC	428	
620 Db	GCNCTTTGGATGGTGGGACGCTCTATTACAGGCAGAGATTTTTNT	TGTGGCCCTTTC	561	
429 QY	CAAAAGGCAAAATCAACGAGGTGCTGAAATCTTTGGCTGATCTTTT	TAAAGCATATGCAGT	488	
560 Db	CAAAAGAAACAAATCAACGAGGTGCTGAAATCTTTGGCTGATCTTTT	TAAAGCATATGCAGT	501	

```

QY 489 CTCACAGTCCAGTGGCAGATGGGTTGCAATTTGAAGAGTTTCTGCAGATGGCTGGGCC 548
DB 500 CTCACAGTCCAGTGGCAGATGGGTTGCAATTTGAAGAGTTTCTGCAGATGGCTGGGCC 441
QY 549 TAACTGATCGCAATTTGGTCTAGTGAATCTGACAGAGGCCCTTAAAGATCATGCAACA 608
DB 440 CAACTGATCGCAATTTGGTCTAGTGAATCTGACAGAGGCCCTTAAAGATCATGCAACA 381
QY 609 GATGAGTCCAGCAGCTACGACAACTCAGCTGCTGATGACATAGCAGCAAACTGTAT 668
DB 380 GATGAGTCCAGCAGCTACGACAACTCAGCTGCTGATGACATAGCAGCAAACTGTAT 321
QY 669 ATATCTAAATATCCCAACAAAGGCGCAGTCTTGTGTCACCAACCCCGGGAAGTATCC 728
DB 320 ATATCTAAATATCCCAACAAAGGCGCAGTCTTGTGTCACCAACCCCGGGAAGTATCC 261
QY 729 AGAAGTCCAAAGGTTTATGAGAACTCAAGGACCATATGCTGATCCCGTGCATGCTC 788
DB 260 AGAAGTCCAAAGGTTTATGAGAACTCAAGGACCATATGCTGATCCCGTGCATGCTC 201
QY 789 TGAATGAAAGGTTGAGTGGGCTGCTCACCTGCTGCTCAGTTTTTAATTAACGAAGGT 848
DB 200 GGAATGAAAGGTTGAGTGGGCTGCTCACCTGCTGCTCAGTTTTTAATTAACGAAGGT 141
QY 849 AGACTCTCTGA 858
DB 140 AGACTCTCTGA 131

RESULT 13
LOCUS AY406329 455 bp DNA linear GSS 15-DEC-2003
DEFINITION Homo sapiens DDH1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY406329
VERSION AY406329.1 GI:39762303
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 455)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 455)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
location/Qualifiers
1..455
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>455
/gene="DDH1"
/locus_tag="HGM2521"

gene
Query Match 52.8%; Score 453.4; DB 29; Length 455;
Best Local Similarity 99.8%; Pred. No. 1.8e-74;
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORIGIN
1..455
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>455
/gene="DDH1"
/locus_tag="HGM2521"

```

```

QY 404 GCAGAGAAATTTTGTGGGCTTTCCAAAGACAAATCCAGAGTGCCTGAATCTTGG 463
DB 1 GCAGAGAAATTTTGTGGGCTTTCCAAAGACAAATCCAGAGTGCCTGAATCTTGG 60
QY 464 CTGATACTTTTAAAGACTATGCAAGTCTCCACAGTGCAGTGGCAGATGGTTGCAATTTGA 523
DB 61 CTGATACTTTTAAAGACTATGCAAGTCTCCACAGTGCAGTGGCAGATGGTTGCAATTTGA 120
QY 524 AGAGTTTCTGCAGCATGGCTGGGCTTAACCTGATGCAATTTGGTCTAGTGAATCTGCAC 583
DB 121 AGAGTTTCTGCAGCATGGCTGGGCTTAACCTGATGCAATTTGGTCTAGTGAATCTGCAC 180
QY 584 AGAAGGCCCTTAAAGATCATGCAACAGATGAGTGCACACCGCTACGACAAATCATCTGTGC 643
DB 181 AGAAGGCCCTTAAAGATCATGCAACAGATGAGTGCACACCGCTACGACAAATCATCTGTGC 240
QY 644 CTGATGACATAGCAGCAAACTGTATATCTATAATATCCCAACAAAGGCGCAGTCTTGC 703
DB 241 CTGATGACATAGCAGCAAACTGTATATCTATAATATCCCAACAAAGGCGCAGTCTTGC 300
QY 704 TGCAACCGAACCCCGGAAGAGTATCCAGAAAGTGCAGAAAGTTTATGAGAAATCTGAAGGACC 763
DB 301 TGCAACCGAACCCCGGAAGAGTATCCAGAAAGTGCAGAAAGTTTATGAGAAATCTGAAGGACC 360
QY 764 ATATGCTGATCCCGTGCAGCATGCTGAACTGGAAAAGTGGATGGCTGCTCACTGCT 823
DB 361 ATATGCTGATCCCGTGCAGCATGCTGAACTGGAAAAGTGGATGGCTGCTCACTGCT 420
QY 824 GCTCAGTTTAAATTAACGAAGGTAGACTCTCTGA 858
DB 421 GCTCAGTTTAAATTAACGAAGGTAGACTCTCTGA 455

```

```

RESULT 14
LOCUS AY406330 455 bp DNA linear GSS 15-DEC-2003
DEFINITION Pan troglodytes DDH1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY406330
VERSION AY406330.1 GI:39762304
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 455)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 455)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
location/Qualifiers
1..455
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>455
/gene="DDH1"
/locus_tag="HGM2521"

gene

```


OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 16:07:05 ; Search time 97 Seconds
(without alignments)
4908.741 Million cell updates/sec

Title: US-09-889-7338-1

Perfect score: 858

Sequence: 1 atggcggcctggccacc.....acaagaagtagactcctga 858

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
 - 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
 - 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
 - 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
 - 5: /cgn2_6/ptodata/2/ina/PCUTUS.COMB.seq:*
 - 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	185.6	21.6	1397	4	US-09-023-655-37
2	81.2	9.5	239	4	US-09-702-705-404
3	81.2	9.5	239	4	US-09-736-457-404
4	81.2	9.5	239	4	US-09-614-124B-404
5	81.2	9.5	239	4	US-09-671-325-404
6	81.2	9.5	239	4	US-09-589-184-404
7	69	8.0	526	4	US-09-621-976-12670
8	55.6	6.5	2598	4	US-09-252-991A-2768
9	55.6	6.5	2790	4	US-09-252-991A-2864
10	52.2	6.1	1288	1	US-08-440-856A-9
11	49.8	5.8	1257	4	US-09-252-991A-1225
12	49.8	5.8	2280	4	US-09-252-991A-1270
13	49.8	5.8	2616	4	US-09-252-991A-1336
14	49.6	5.8	7218	1	US-08-232-463-14
15	49.4	5.8	666	6	5472691-4
16	49.4	5.8	669	3	US-08-556-965-1
17	49.4	5.8	723	6	5472691-7
18	49.4	5.8	1239	4	US-09-252-991A-6622
19	49.4	5.8	1389	2	US-08-023-980B-3
20	49.4	5.8	1389	2	US-08-486-953A-3
21	49.4	5.8	1396	6	5472691-1
22	49.4	5.8	1449	4	US-09-252-991A-6642
23	49.4	5.8	1647	4	US-09-252-991A-6576
24	49.4	5.8	1731	4	US-08-252-991A-6675
25	49.4	5.8	10079	2	US-08-476-866-20
26	48.2	5.6	1389	4	US-09-252-991A-4243
27	48.2	5.6	2598	4	US-09-252-991A-4323

C	28	48.2	5.6	2874	4	US-09-252-991A-4533	Sequence 4533, Ap
C	29	47.6	5.5	43280	2	US-08-804-227C-1	Sequence 1, Appli
C	30	47.2	5.5	47981	4	US-09-679-279-1	Sequence 1, Appli
C	31	47	5.5	1074	4	US-09-252-991A-5833	Sequence 5833, Ap
C	32	47	5.5	1275	4	US-09-252-991A-5813	Sequence 5813, Ap
C	33	47	5.5	1554	4	US-09-252-991A-5777	Sequence 5777, Ap
C	34	46.6	5.4	1187	1	US-08-440-856A-2	Sequence 2, Appli
C	35	46.4	5.4	519	4	US-09-252-991A-16355	Sequence 16355, A
C	36	46.4	5.4	609	4	US-09-252-991A-16464	Sequence 16464, A
C	37	46.4	5.4	4257	2	US-08-690-473-1	Sequence 1, Appli
C	38	46.4	5.4	4257	3	US-09-259-821A-1	Sequence 1, Appli
C	39	46.4	5.4	4257	3	US-08-843-859-1	Sequence 1, Appli
C	40	46.4	5.4	12001	1	US-08-458-568A-11	Sequence 11, Appli
C	41	46.2	5.4	1926	4	US-09-249-585A-4	Sequence 4, Appli
C	42	46.2	5.4	1931	2	US-09-130-114-2	Sequence 2, Appli
C	43	46.2	5.4	11220	3	US-09-105-537-32	Sequence 32, Appli
C	44	46.2	5.4	15664	1	US-08-402-282-3	Sequence 3, Appli
C	45	46.2	5.4	15664	1	US-08-508-004-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-023-655-37

; Sequence 37, Application US/09023655

; Patent No. 6607879

; GENERAL INFORMATION:

; APPLICANT: Cocks, Benjamin G.

; APPLICANT: Susan G. Stuart

; APPLICANT: Jeffrey J. Sealhamer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE

; NUMBER OF SEQUENCES: 1508

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/023,655

; FILING DATE: HEREWITH

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Zeller, Karen J.

; REGISTRATION NUMBER: 37,071

; REFERENCE/DOCKET NUMBER: PA-0001 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 855-0555

; TELEFAX: (650) 845-4166

; INFORMATION FOR SEQ ID NO: 37:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1397 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: HMC1NOT01

; CLONE: 003437

US-09-023-655-37

Query Match

21.6%; Score 185.6; DB 4; Length 1397;

Best Local Similarity 57.4%; Pred. No. 4.7e-35;
Matches 372; Conservative 0; Mismatches 274; Indels 2; Gaps 2;
QY 29 TCGCGCGCCACCCACCGCGTGGTGGCGCGCTACCCGAGTGGTGGCGAGCACGCGC 88
Db 357 TGGGCGCGTCTCCCATGATCCCGGGAGTCCAGAGAGCGTGGCGTGGGGGAAG 416
QY 89 TGAGAAGCGCCAAAGCGCGAGAGTGGAGCTCGCCCGCGCGAAACGGCAGCACAGCTCT 148
Db 417 GTGCGGGGCTGGCTTCCGCTCTGGATCTGGCCAAAGCTCAAAGGGAGCACGGGTGC 476
QY 149 AGCTGGCGGTGTGGCGAGCAAGCTGGGGCTGTGAGGTGGTGGAGTGGCGCGCGAGAGA 208
Db 477 TGGGAGGTAACTGAGCAACAGCTGGGGCTACAGCTGTAGAAGTGCACCTGAGGAGT 536
QY 209 GCTTTCCGAGCTGCTCTTCGTTGGAGGAGCTGGCGCGTGTGTGGAGAGACGGCCCTCA 268
Db 537 CATTCGCGTGGACCGCTGTGTGGGACACGGCCGTGATCAAGGGGACACGGCCCTAA 596
QY 269 TCACCGGACCGCGCGCGAGCGAGCGAGAGAGGTGATGATGATGAAGAGCATTAG 328
Db 597 TCACGCGCGCTGGAGCGCCCGCTCGTAGGCCAGAGGTGATGGAGTCCGCAAGCCCTGC 656
QY 329 AAAAACTTCAGCTCAATATAGTAGAGTCAAGATGAAATGCAACTTTAGATGGCGAG 388
Db 657 AAGACTGGGGTCCGAATTTGGGAATAGAGACAGAAACGGAGCTGGTGGCACTG 716
QY 389 ATGTTTATTCAGCGAGAGAAATTTTGGGGCTTTCCAAAAGGACAAATCAACGAG 448
Db 717 AGCTTCTCTTACCGCGCGGAGTTTTTCGTAGGGCTCTCCAAATGGACCAATCACGAG 776
QY 449 GTGCT-GAAATCTGGCTGATATTTTAAAGACTATGAGTCCACAGTCCAGTGGCA 507
Db 777 GCGTGGAGATCGTGGGACACGTTCCGGGACTTCCGGCGTCTCCACTGTGCCAGTCTCG 836
QY 508 GATGGTGTCAATTTGAAGAGTTTTCGACAGATGCTGGCGCTTAACCTGATCGCAATTTGG 567
Db 837 GGTCCCTCCACCTGGCGGTCTCTCGCGCATGGGGGACCTCGCACTGTTGTGGCAGGC 896
QY 568 TGTAGTGAATCTGCAGAGAGCCCTTAAGATCATGCAACAGATGAGTGCACCGGTAC 627
Db 897 ACAGCGAGCTGGCCAAAGGCTTCCGGGCAATGGCAGTGGTGCACAGATCAACCCATAT 956
QY 628 GACAACTCA-CTGTGCTGTATGATACATAGCAGCAAACTGTATATCT 674
Db 957 GCTCCCTGACCCCTCCAGATGACGAGCTGCTGACTGTCTCTTCT 1004

RESULT 2
US-09-702-705-404/c
; Sequence 404, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien

US-09-702-705-404
Query Match 9.5%; Score 81.2; DB 4; Length 239;
Best Local Similarity 58.8%; Pred. No. 1.9e-10;
Matches 140; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 185 TGGTGGAGTGGCGCGCGAGAGCCTTCCGAGTGGCTTTCGTGGAGGAGTGGCGCG 244
Db 238 TGCTAGAACTGCCACCTGAGGAGTCAATTCGCGGTGGGACCGCTTGGCGACACGCGCG 179
QY 245 TGGTGTGGAGGAGAGCGCCCTCATCCCGACCGCGCGCGCGCGAGCGAGGAGG 304
Db 178 TGATCCAAAGGGGACACGGCCCTAATCACGGGCGCTTGGAGCCCGCTCGTAGCCAGAG 119
QY 305 TTGACATGATGAAGAAGCATTTAGAAAACTTCAGCTCAATATAGTAGAGATGAAGAATG 364
Db 118 TCGATGGAGTCCGCAAGCCCTGCAAGACTGGGGCTCCGAATTTGGAAATAGGAGAGCG 59
QY 365 AAATGCAACTTTAGATGGCGGAGATGTTTTTATTCACAGCGAGAGATTTTGTGGG 422
Db 58 AGAACGCGAGCTGGATGGCACTGCTCTTCTTACCGCGCGGAGTTTTTCGTAGG 1
RESULT 3
US-09-736-457-404/c
; Sequence 404, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien

US-09-736-457-404
Query Match 9.5%; Score 81.2; DB 4; Length 239;
Best Local Similarity 58.8%; Pred. No. 1.9e-10;
Matches 140; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 185 TGGTGGAGTGGCGCGCGAGAGCCTTCCGAGTGGCTTTCGTGGAGGAGTGGCGCG 244
Db 238 TGCTAGAACTGCCACCTGAGGAGTCAATTCGCGGTGGGACCGCTTGGCGACACGCGCG 179
QY 245 TGGTGTGGAGGAGAGCGCCCTCATCCCGACCGCGCGCGCGAGCGAGGAGGAGG 304
Db 178 TGATCCAAAGGGGACACGGCCCTAATCACGGGCGCTTGGAGCCCGCTCGTAGCCAGAG 119
QY 305 TTGACATGATGAAGAAGCATTTAGAAAACTTCAGCTCAATATAGTAGAGATGAAGAATG 364
Db 118 TCGATGGAGTCCGCAAGCCCTGCAAGACTGGGGCTCCGAATTTGGAAATAGGAGAGCG 59
QY 365 AAATGCAACTTTAGATGGCGGAGATGTTTTTATTCACAGCGAGAGATTTTGTGGG 422
Db 58 AGAACGCGAGCTGGATGGCACTGCTCTTCTTACCGCGCGGAGTTTTTCGTAGG 1
RESULT 4

US-09-614-124B-404/C
; Sequence 404, Application US/09614124B
; Patent No. 8630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-404

Query Match
Best Local Similarity 9.5%; Score 81.2; DB 4; Length 239;
Matches 140; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 185 TGGTGGAGCTCCCGCGGACGAGAGCCTTCGCGACTCGCTTCGTGGAGGACGTCGCGG 244
DB 238 TGCTAGAACTCCACCTGAGGAGTCATTGCGCTGGGACCGCTGCTTGGCGACACGGCGG 179
QY 245 TGGTGTGGAGGAGACGGCCCTCATCCCGACCGGGCGCGCGGAGGAGGAGG 304
DB 178 TGATCCAAAGGGGACACGGCCCTTAATCAGCGGCGCTGGAGCCCGCTGCTAGGCGAGG 119
QY 305 TTGACATGATGAAGAAGCATTAGAAAACTTCAGCTCAATATATAGTAGAGATGAAGATG 364
DB 118 TCGATGGAGTCCGAAAGCCCTGCAAGACCTGGGGCTCCGAATTTGTGAAATAGGAGAGC 59
QY 365 AAAATGCAACTTTAGATGGCGGAGATGTTTTATTTCACAGGACAGAAATTTTTTGTGG 422
DB 58 AGAACGGACGCTGGATGGCACTGACGCTTCTTTCACCGCGCGGAGTTTTTCGTAGG 1

RESULT 6
US-09-589-184-404/C
; Sequence 404, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; CURRENT FILING DATE: 2000-06-05
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-589-184-404

Query Match
Best Local Similarity 9.5%; Score 81.2; DB 4; Length 239;
Matches 140; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 185 TGGTGGAGCTCCCGCGGACGAGAGCCTTCGCGACTGGTTCGTGGAGGACGTCGCGG 244
DB 238 TGCTAGAACTCCACCTGAGGAGTCATTGCGCTGGGACCGCTGCTTGGCGACACGGCGG 179
QY 245 TGGTGTGGAGGAGACGGCCCTCATCCCGACCGGGCGCGCGGAGGAGGAGG 304
DB 178 TGATCCAAAGGGGACACGGCCCTTAATCAGCGGCGCTGGAGCCCGCTGCTAGGCGAGG 119
QY 305 TTGACATGATGAAGAAGCATTAGAAAACTTCAGCTCAATATATAGTAGAGATGAAGATG 364
DB 118 TCGATGGAGTCCGAAAGCCCTGCAAGACCTGGGGCTCCGAATTTGTGAAATAGGAGAGC 59
QY 365 AAAATGCAACTTTAGATGGCGGAGATGTTTTATTTCACAGGACAGAAATTTTTTGTGG 422
DB 58 AGAACGGACGCTGGATGGCACTGACGCTTCTTTCACCGCGCGGAGTTTTTCGTAGG 1

RESULT 7
US-09-621-976-12670
; Sequence 12670, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:

US-09-614-124B-404/C
; Sequence 404, Application US/09614124B
; Patent No. 8630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-404

Query Match
Best Local Similarity 9.5%; Score 81.2; DB 4; Length 239;
Matches 140; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 185 TGGTGGAGCTCCCGCGGACGAGAGCCTTCGCGACTCGCTTCGTGGAGGACGTCGCGG 244
DB 238 TGCTAGAACTCCACCTGAGGAGTCATTGCGCTGGGACCGCTGCTTGGCGACACGGCGG 179
QY 245 TGGTGTGGAGGAGACGGCCCTCATCCCGACCGGGCGCGCGGAGGAGGAGG 304
DB 178 TGATCCAAAGGGGACACGGCCCTTAATCAGCGGCGCTGGAGCCCGCTGCTAGGCGAGG 119
QY 305 TTGACATGATGAAGAAGCATTAGAAAACTTCAGCTCAATATATAGTAGAGATGAAGATG 364
DB 118 TCGATGGAGTCCGAAAGCCCTGCAAGACCTGGGGCTCCGAATTTGTGAAATAGGAGAGC 59
QY 365 AAAATGCAACTTTAGATGGCGGAGATGTTTTATTTCACAGGACAGAAATTTTTTGTGG 422
DB 58 AGAACGGACGCTGGATGGCACTGACGCTTCTTTCACCGCGCGGAGTTTTTCGTAGG 1

RESULT 5
US-09-671-325-404/C
; Sequence 404, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-404

APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621.976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 12670
LENGTH: 526
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-12670

Query Match 8.0%; Score 69; DB 4; Length 526;
Best Local Similarity 52.2%; Pred. No. 1.9e-07;
Matches 153; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 88 CTGAGAGCCGCAAGGCGAGGAGTGGAGCTGCCCGCGGGAACGCGACGACGCTC 147
DB 234 CTAGAATAATCGAGCGGTGGAGGGGAGTCTGTGTGATGGGAGCGCGGGTG 293
QY 148 TACGTGGCGTGTGGGAGCAAGCTGGGCTGCAGTGTGGAGTGGCGGCGGACGAG 207
DB 294 CTGGAGGTAACTGAGCAACGACTGGGCTACAGCTGTAGAACTGCCACCTGAGGAG 353
QY 208 AGCCTTCGGACTCGCTTCTGTGAGGAGTGGCGTGGTCCGAGGAGCGGCTC 267
DB 354 TCATTGCCGCTGGGACCGCTCTTGGGACACGCGCGGTGATCAAGGGGACAGGCCCTA 413
QY 268 ATCAACCCACCGGCGCGCGGAGGAGGAGGAGTGCATGATGAAAGAGCATTA 327
DB 414 ATCAGCGCGCTCGAGCGCGCTCGTAGGCCAGAGTGCATGAGTCCGCAAGCCCTG 473
QY 328 GAAAACTTCAGCTCAATATAGTAGAGTGAAGATGAAATGCAATTTAGA 380
DB 474 CAAGACCTGGGCGCTCCGAATTTGGAAATAGGAGACGAGACGCGACGCTGGA 526

RESULT 8

US-09-252-991A-2768
Sequence 2768, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR FILING DATE: 1998-02-18
PRIOR FILING DATE: 1998-02-18
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 2768
LENGTH: 2598
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2768

Query Match 6.5%; Score 55.6; DB 4; Length 2598;
Best Local Similarity 49.2%; Pred. No. 0.00052;
Matches 176; Conservative 0; Mismatches 179; Indels 3; Gaps 1;

QY 83 ACGCGCTGAGAACGCCAAGGCGGAGGAGTGGAGCTGCCCGCGGGAACGCGACGACC 142
DB 357 ACGGCTTGACCTCCAGCCACCTCGGCAAGCGGACTACGCCAAGGCCCTGGAGCAGCACA 416
QY 143 AGCTTACGTGGCGGTGTGGGACACAGCTGGGCTCAGGTGTGGAGTGGCGGCG 202
DB 417 ACGCCTACATCCGCGCTTGCAGACC---TGCAGCTGACATCACCTGCTGCGCGCG 473

QY 203 ACGAGAGCTTCCGAGCTGCGTCTTGTGAGGAGAGTGGCGGTGTGTGCGAGGAGCGG 262
DB 474 ACGAACGCTTCCCGAGCTCGGTGTGTGTCGAGGAGCCCGGTGTGTGACCTCGCGCTCG 533
QY 263 CCTCATCACCGACCGCGCGGCGCGAGGAGGAGTTCACATGATGAAAGAG 322
DB 534 CCATCATCACCGCGCGCGCGCGGCGGCGGAGTTCGCGCGGCGGAGACCGAGATCATCGAGGAAA 593
QY 323 CATTAGAAAACCTTCAGCTCAATATAGTAGAGTGAAGATGAAATGCAACTTTAGATG 382
DB 594 CCGTGCAGCGCTTCTATCCGGGCAAGGTTCGAGCGCATCGAGGACCCCGGCGACG 653
QY 383 GCGGAGATGTTTATTACAGGACAGAGATTTTGTGGGCTTTCAAAAGGACAAA 440
DB 654 CCGGCGACATCATGATGTCGCGGACCACTTCTATACCGGCAATCGGCCCGCACCAA 711

RESULT 9

US-09-252-991A-2864
Sequence 2864, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR FILING DATE: 1998-02-18
PRIOR FILING DATE: 1998-02-18
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 2864
LENGTH: 2790
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2864

Query Match 6.5%; Score 55.6; DB 4; Length 2790;
Best Local Similarity 49.2%; Pred. No. 0.00053;
Matches 176; Conservative 0; Mismatches 179; Indels 3; Gaps 1;

QY 83 ACGCGCTGAGAACGCCAAGGCGGAGGAGTGGAGCTGCCCGCGGGAACGCGACGACC 142
DB 2072 ACGGCTGACCTCCAGCCACCTTCGCAAGCGCGGACTACGCCAAGGCCCTTGGAGCAGCACA 2131
QY 143 AGCTTACGTGGCGGTGTGGCAGCAAGCTGGGCTGCAGGTGTGGAGTGGCGGCG 202
DB 2132 ACGCTTACATCCGCGCTTGCAGACC---TGCAGCTGAGACATCACCTGCTGCCGCGG 2188
QY 203 ACGAGAGCTTCCGAGCTCGCTTTCGTGGAGGAGCTGGCGCTGTGTGTCGAGGAGACGG 262
DB 2189 ACGAACGCTTCCCGACTCGGTGTTCGTGAGGAGCCCGGTGTCTGCACCTCGCGTGG 2248
QY 263 CCTCATCACCGACCGCGGCGCGGCGGAGGAGGAGTTCACATGATGAAGAG 322
DB 2249 CCATCATCACCGCGCGCGCGCGGCGGAGTTCGCGCGGCGGAGACCGAGATCATCGAGGAAA 2308
QY 323 CATTAGAAAACCTTCAGCTCAATATAGTAGAGTGAAGATGAAATGCAACTTTAGATG 382
DB 2309 CCGTGCAGCGCTTCTATCCGGGCAAGGTTCGAGCGCATCGAGGACCCCGGCGCGTGAAG 2368
QY 383 GCGGAGATGTTTATTACAGGACAGAGATTTTGTGGGCTTTCAAAAGGACAAA 440
DB 2369 CCGGCGACATCATGATGTCGCGGACCACTTCTATACGCGGCAATCGGCCCGCACCA 2426

RESULT 10

US-08-440-856A-9
Sequence 9, Application US/08440856A
Patent No. 5750873

GENERAL INFORMATION:
APPLICANT: DELLAPORTA, STEPHEN L.
TITLE OF INVENTION: MATERIALS AND METHODS FOR PRODUCING
TITLE OF INVENTION: PLANTS WITH SINGLE-SEX FLOWERS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FORSTER
STREET: 2000 PENNSYLVANIA AVE. N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

FILING DATE: 15-MAY-1995
CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: MILLMAN, ROBERT A.

REGISTRATION NUMBER: 36,217

REFERENCE/DOCKET NUMBER: 05463-20001.00

TELEPHONE: (202) 887-1517

TELEFAX: (202) 887-0763

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 1288 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-440-856A-9

Query Match 6.1%; Score 52.2; DB 1; Length 1288;

Best Local Similarity 41.8%; Pred. No. 0.0026;

Matches 105; Conservative 29; Mismatches 117; Indels 0; Gaps 0;

QY 11 TCGGCCACCCCTCGGCTTCGGCGGGCCACCCACCGCGTGTGCGGGGCTACCCGAGT 70

DB 343 TCGGACATCGACGAGCGCGGGGAGCGCTGCGCGCGCTGCGCGCGCGTCTC 402

QY 71 CGCTCTGCCAGCAGCGCTGAGAACGCCAACGCGGAGGAGTGGACGTCGCCCGCGG 130

DB 403 GSITCGTGGTGGAGTGTCTGCGGAGGAGTGTGCGGAGGAGTGTGCGGAGGAGT 462

QY 131 AACGGGACACGAGTGTCTGCGGAGGAGTGTGCGGAGGAGTGTGCGGAGGAGT 190

DB 463 YGTSGCGSRYGCGGCGCTGAGTGTCTGCGGAGGAGTGTGCGGAGGAGTGTGCGG 522

QY 191 AGCTGCGCGGCGAGGAGTGTGCGGAGGAGTGTGCGGAGGAGTGTGCGGAGGAGT 250

DB 523 AGACGCGCGGCGAGGAGTGTGCGGAGGAGTGTGCGGAGGAGTGTGCGGAGGAGT 582

QY 251 GCGAGGAGGAGC 261

DB 583 GCGTCAACGCG 593

RESULT 11

US-09-252-991A-1225

Sequence 1225; Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 1225

LENGTH: 1257

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-1225

Query Match 5.8%; Score 49.8; DB 4; Length 1257;

Best Local Similarity 48.7%; Pred. No. 0.0097;

Matches 135; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 2 TGGCCGGCTCGGCGACCCCTTCGCGCGGGCCACCCACCGCGTGTGCGGGCGC 61

DB 717 TGGTCGACATCGGCAACCCGCTGCGCGACCGCAGGACCTGAGCACCGTCTCGACGTGA 776

QY 62 TACCCGAGTCTGCGCAGCAGCGCTGAGAGCGCAAGCGCGGAGGAGTGGACGTG 121

DB 777 TGAACCACTGATCAGCTGATGACCCGGAAGAACTCCGAAGCCAGCTGGCCACGCTCA 836

QY 122 CCCGCGGCGGACGAGCAGCAGCTTACGTTGGCGTGTGGCGAGCAAGCTGGGGTGC 181

DB 837 AGCCCGCGGAGCTGCTGATCCAGCCGCGCTGTCGGCTACGGCACCCAGCTTCGCGC 896

QY 182 AGTGTGTGAGCTCCCGCGCGCAGCAGGACCTTCGAGTCTGCTGTGGAGACGTGG 241

DB 897 GCGTGGCGCAACTGATCGAGCGCGGTACCGCGCACCGAGTCTGCGCGCGCGCTCG 956

QY 242 CCGTGTGTGCGGAGGAGCGGCTCATCACCGGACC 278

DB 957 CCGAGTTGCGCAAGCCGAGGACCTCAACAGCGAAGC 993

RESULT 12

US-09-252-991A-1270

Sequence 1270; Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 1270

LENGTH: 2280

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-1270

Query Match 5.8%; Score 49.8; DB 4; Length 2280;

Best Local Similarity 48.7%; Pred. No. 0.012; Indels 0; Gaps 0;

Matches 135; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 2 TGGCCGGCTCGGCGACCCCTTCGCGCGGGCCACCCACCGCGTGTGCGGGCGC 61

DB 779 TGGTCGACATCGGCAACCCGCTGCGCGACCGCAAGGACCTGAGCACCGTCTCGACGTGA 838

QY 62 TACCCGAGTCTGCGCAGCAGCGCTGAGAGCGCAAGCGCGGAGGAGTGGACGTG 121

DB 839 TGAACCACTGATCAGCTGATGACCCGGAAGAACTCCGAAGCCAGCTGGCCACGCTCA 898

QY 122 CCCGCGGCGGACGAGCAGCAGCTTACGTTGGCGTGTGGCGAGCAAGCTGGGGTGC 181

```

899  AGCCGCGGACGTGCTGATCCAGCGCCGCTCTCGGCTACGCGACACCGACTTCGGCC 958
182  AGGTGTTGAGCTGCCGGCCGACGAGAGCCTTCCGAGATGCTCTTCGTGGAGGACGTGG 241
959  GCGTGGCCGAACTGATTCGACGCGCGGTACCGCGCCACCAACGCTGCTCGCGCGCCTCG 1018
242  CCGTGTGTGCGGAGGAGCGGCGCTCATCACCGACC 278
1019  CCGAGTTGCGCAAGCCGAGGACCTCAACAGCGAACC 1055

RESULT 13
US-09-252-991A-1336/c
; Sequence 1336, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1336
; LENGTH: 2616
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1336

```

Query Match	5.8%;	Score 49.8;	DB 4;	Length 2616;
Best Local Similarity	48.7%;	Pred. No. 0.012;		
Matches 135;	Conservative 0;	Mismatches 142;	Indels 0;	Gaps 0;
QY	2	TGGCCGGCCTTCGGCCACCCCTTCGCGCGGCCACCAACGCGTGTGGCGGCGC	61	
Db	1517	TGCTGACATCGGCACACCCGCTCGGACCGCAAGGACCTGAGCAACGCTCTCGACGTGA	1458	
QY	62	TACCGAGTCGCTCTGCCACACCGCTGAGAGAGCCCAAGGCGAGGAGGTGGACGTGC	121	
Db	1457	TGAACCAAGTCGATCAACCTGATACCCGGAGAACTCCGAGCCAGCTGGCCACGCTCA	1398	
QY	122	CCGCGCGGAACGGCAGCACCAAGCTCTAAGTGGCGGTGCTGGGCAGCAAGCTGGGCGCTGC	181	
Db	1397	AGCCCGCGACGTGCTGATCCAGCCGCGGTGCTCCGGCTACGGCACCACCGACTTCGGCC	1338	
QY	182	AGTGGTGGAGCTCCCGCGCAGAGAGCCTTCGGACTGCGTCTTCGTGGAGGACGTGG	241	
Db	1337	CGGTGCCCAACTGATCGACCGCGTACCGCGCCACCAGGTGCTGCGCGCGCGCTCG	1278	
QY	242	CCGTGGTGTGCGGAGACGGCCCTCATCACCGAAC	278	
Db	1277	CCGAGTTGCGCAAGCCGAAGGACCTCAACAGCGAAGC	1241	

RESULT 14
US-08-232-463-14/c
; Sequence 14, Application US/06232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/125,744
; FILING DATE: 24-SP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 897,624
; FILING DATE: 12-JUN-1992
; APPLICATION NUMBER: 576,114
; FILING DATE: 27-aug-1990
; APPLICATION NUMBER: 902,596
; FILING DATE: 02-sep-1986
; SEQ ID NO:4:
; LENGTH: 666
5472691-4

Query Match      5.8%; Score 49.4; DB 6; Length 666;
Best Local Similarity 50.2%; Pred. No. 0.0098;
Matches 122; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 23 CGCCTTCGGCGGCGCACCGCGTGGTGGGGCGCTACCGAGTCGCTCTGCCAGC 82
Db 593 CGGGCCCCGACACGCCACCGCCAGCGAGCGCCAGCCCGCGCGCTTCCCTTCC 534
QY 83 ACAGCGCTGAGAGCGCCCAAGGGCGAGAGGTGGACGTGCCCGCGCGGAACGGCAGCACC 142
Db 533 ACAGCGCTGAGAGCGCCCAAGGGCGAGAGGTGGACGTGCCCGCGCGGAACGGCAGCACC 474
QY 143 AGCTTACGTGGGCGTGGCGAGCAAGCTGGGGCTGAGTGGTGGAGCTGCCGGCGG 202
Db 473 CGGCCACCATGATGAGTGGCGGGCGCGCGAGCGCGCCAGCGCGCGGTACCTCCAG 414
QY 203 ACAGAGAGCCTTCGGGACTGCGTCTTCGTGGAGGACGTGGCCGCTGGTGTGCGAGGAGACGG 262
Db 413 AGGCTCGCGTGGGAGCGCGGAGTTGCCAGAGTCCGCCGGGTGCTGGGGTGGCGGACG 354
QY 263 CCC 265
Db 353 GCC 351

```

```

Search completed: June 7, 2004, 20:34:32
Job time : 100 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 19:33:33 ; Search time 417 seconds
(without alignments)
9386.553 Million cell updates/sec

Title: US-09-889-733B-1

Perfect score: 858
Sequence: 1 atggcggcctggcacc.....acaagaagtagactctctga 858

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2995936 seqs, 228098010 residues

Total number of hits satisfying chosen parameters: 5991872

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/prodata/1/pubpna/PCT05_PUBCOMB.seq.*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq.*
11: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq.*
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq.*
14: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	851.6	99.3	1633	17	US-10-450-826-9
2	851.6	99.3	4208	17	US-10-468-125-22
3	715.6	83.4	3008	15	US-10-205-219-83
4	705.4	82.2	3278	15	US-10-252-157-328
5	630.4	73.5	3651	9	US-09-925-300-512
6	197.8	23.1	770	9	US-09-910-943-9
7	185.6	21.6	1397	17	US-10-641-643-37
8	156	18.2	332	9	US-09-783-590-5309
9	148.4	17.3	472	10	US-09-918-995-9117
10	148	17.2	83946	17	US-10-450-826-8
11	146	17.0	452	10	US-09-918-995-14871
12	134.2	15.6	403	10	US-09-918-995-19254
13	119.4	13.9	3189	10	US-09-814-353-20553
14	81.2	9.5	239	9	US-09-736-457-404

c 15 81.2 9.5 239 9 US-09-902-941-404 Sequence 404, App
c 16 81.2 9.5 239 9 US-09-849-626-404 Sequence 404, App
c 17 81.2 9.5 239 10 US-09-476-300-404 Sequence 404, App
c 18 81.2 9.5 239 13 US-10-283-017-404 Sequence 404, App
c 19 81.2 9.5 239 15 US-10-017-754-404 Sequence 404, App
c 20 81.2 9.5 239 15 US-10-113-872-404 Sequence 404, App
c 21 65.2 7.6 774 15 US-10-156-761-1679 Sequence 1679, App
c 22 65.2 7.6 9025608 15 US-10-156-761-1 Sequence 1, Appli
c 23 60.2 7.0 975 13 US-10-107-431-14 Sequence 14, Appli
c 24 60.2 7.0 45055 13 US-10-107-431-277 Sequence 277, App
c 25 60 60 10 US-09-908-975-9030 Sequence 9030, App
c 26 56.8 6.6 1647 13 US-10-425-114-28074 Sequence 28074, A
c 27 55 6.4 3729 13 US-10-671-403-86 Sequence 86, Appli
c 28 55 6.4 3729 13 US-10-671-419-86 Sequence 86, Appli
c 29 55 6.4 3729 13 US-10-670-844-86 Sequence 86, Appli
c 30 55 6.4 3729 13 US-10-671-134-86 Sequence 86, Appli
c 31 55 6.4 3729 13 US-10-673-098-86 Sequence 86, Appli
c 32 55 6.4 3729 17 US-10-672-638-86 Sequence 86, Appli
c 33 55 6.4 3729 17 US-10-673-127-86 Sequence 86, Appli
c 34 52.4 6.1 3012 9 US-09-738-626-2176 Sequence 2176, App
c 35 52.4 6.1 3309400 9 US-09-738-626-1 Sequence 1, Appli
c 36 52 6.1 903 9 US-09-815-242-4130 Sequence 4130, App
c 37 52 6.1 903 13 US-10-282-122A-7433 Sequence 7433, App
c 38 52 6.1 2385 16 US-10-369-493-31665 Sequence 31665, A
c 39 51 5.9 2073 12 US-09-758-759-158 Sequence 158, App
c 40 51 5.9 7788 15 US-10-329-079-8 Sequence 8, Appli
c 41 51 5.9 37360 15 US-10-329-079-6 Sequence 6, Appli
c 42 51 5.9 109519 12 US-09-758-759-1 Sequence 1, Appli
c 43 50.8 5.9 585 15 US-10-156-761-6221 Sequence 6221, App
c 44 50.8 5.9 1749 13 US-10-282-122A-25460 Sequence 25460, A
c 45 50.8 5.9 9025608 15 US-10-156-761-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-10-450-826-9
; Sequence 9, Application US/10450826
; Publication No. US20040101818A1
; GENERAL INFORMATION:
; APPLICANT: Ji, Darren
; APPLICANT: Axelrod, Douglas W.
; APPLICANT: Cook, Jonathan S.
; APPLICANT: Jaiswal, Neelam
; APPLICANT: Eistein, Richard
; APPLICANT: Houghton, Adam
; APPLICANT: Mertz, Lawrence
; TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiation
; FILE REFERENCE: 044921-5039-WO
; CURRENT APPLICATION NUMBER: US/10/450,826
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 60/255,882
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/285,691
; PRIOR FILING DATE: 2001-04-24
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1633
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. AB001915
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1633)
; OTHER INFORMATION: n = a or c or g or t
US-10-450-826-9

Query Match 99.3%; Score 851.6; DB 17; Length 1633;
Best Local Similarity 99.5%; Pred. No. 6.5e-235;
Matches 854; Conservative 0; Mismatches 4; Indels 0; Gaps 0;


```

QY 1 ATGCCGCGCTCGCCACCCCTCCGCTTGGCGGGCCACCCACCGCGTGGTGGCGGG 60
Db 323 ATGCCGCGCTCGCCACCCCTCCGCTTGGCGGGCCACCCACCGCGTGGTGGCGGG 382
QY 61 CTACCCGAGTGGCTCTGCGAGCAGCGCTGAGAGCGCCGAGGCGAGAGTGGCGTC 120
Db 383 CTACCCGAGTGGCTCTGCGAGCAGCGCTGAGAGCGCCGAGGCGAGAGTGGCGTC 442
QY 121 GCCCGCGGAAACCGCAGCAGCAGCTCTAGCTGGGCGCTGCTGGGCGAGCAGCTGGGCGCTG 180
Db 443 GCCCGCGGAAACCGCAGCAGCAGCTCTAGCTGGGCGCTGCTGGGCGAGCAGCTGGGCGCTG 502
QY 181 CAGTGTGTGAGTGGCGGCGGCGAGAGCGCTTCGAGCTTCGAGTGGTGGGAGGAGCTG 240
Db 503 CAGTGTGTGAGTGGCGGCGGCGAGAGCGCTTCGAGCTTCGAGTGGTGGGAGGAGCTG 562
QY 241 GCCGTGTGTGCGAGGAGAGCGGCGCTCATCACCGACCCGCGGCGCGAGCGCGGAGGAG 300
Db 563 GCCGTGTGTGCGAGGAGAGCGGCGCTCATCACCGACCCGCGGCGCGAGCGCGGAGGAG 622
QY 301 GAGTGTGACATGATGAAGAGCATTAGAAAATTCAGCTCAATATAGTAGAGATGAAA 360
Db 623 GAGTGTGACATGATGAAGAGCATTAGAAAATTCAGCTCAATATAGTAGAGATGAAA 682
QY 361 GATGAAATGCAACTTTAGATGGGCGAGATGTTTATTCACAGCGAGAGATTTTGTG 420
Db 683 GATGAAATGCAACTTTAGATGGGCGAGATGTTTATTCACAGCGAGAGATTTTGTG 742
QY 421 GCGCTTTCCAAAAGGACAAATCAACGAGTGGCTGAAATCTTGGCTGATATTTTAAGGAC 480
Db 743 GCGCTTTCCAAAAGGACAAATCAACGAGTGGCTGAAATCTTGGCTGATATTTTAAGGAC 802
QY 481 TATGAGTCTCCAGTGGCGAGTGGGTTGCAATTTGAAGAGTTCCTCAGCATG 540
Db 803 TATGAGTCTCCAGTGGCGAGTGGGTTGCAATTTGAAGAGTTCCTCAGCATG 862
QY 541 GCTGGGCTACCTGATGCGCAATTTGGTCTAGTGAATCTGCACAGAGGCGCTTAAGATC 600
Db 863 GCTGGGCTACCTGATGCGCAATTTGGTCTAGTGAATCTGCACAGAGGCGCTTAAGATC 922
QY 601 ATGCAACAGATGAGTGACCAACCGCTACGAAATCTACTGTGCTGATGACATAGCAGCA 660
Db 923 ATGCAACAGATGAGTGACCAACCGCTACGAAATCTACTGTGCTGATGACATAGCAGCA 982
QY 661 AACTGTATATATCTAATATATATATATATATATATATATATATATATATATATATAT 720
Db 983 AACTGTATATATCTAATATATATATATATATATATATATATATATATATATATATAT 1042
QY 721 GAGTATCCAGAAAGTCAAAAGGTTTATGAGAACTGAAGAACCATATGCTGATCCCGGTG 780
Db 1043 GAGTATCCAGAAAGTCAAAAGGTTTATGAGAACTGAAGAACCATATGCTGATCCCGGTG 1102
QY 781 AGCATGTCTGAACCTGGAAGGTTGGATGGGCTGCTACCTGCTGCTCAGTTTAAATTAAC 840
Db 1103 AGCATGTCTGAACCTGGAAGGTTGGATGGGCTGCTACCTGCTGCTCAGTTTAAATTAAC 1162
QY 841 AAGAAAGTAGACTCTCTGA 858
Db 1163 AAGAAAGTAGACTCTCTGA 1180

```

RESULT 2

```

US-10-468-125-22
; Sequence 22, Application US/10468125
; Publication No. US20040082061A1
; GENERAL INFORMATION:
; APPLICANT: ASTROMOFF, Anna
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BAUGHN, Marian R.
; APPLICANT: DING, Li
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: FORSYTHE, Ian J.

```

```

; APPLICANT: GIETZEN, Kimberly J.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: LU, Yan
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: RING, Huijun Z.
; APPLICANT: SANJANWALA, Madhusudan
; APPLICANT: SWARNAXAR, Anita
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: WARREN, Bridget A.
; APPLICANT: XU, Yuming
; APPLICANT: YUE, Henry
; APPLICANT: ZEBARJADIAN, Yeganeh
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0363 USN
; CURRENT APPLICATION NUMBER: US/10/468.125
; CURRENT FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: PCT/US02/04918
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,643
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/271,332
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/276,767
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: US 60/282,077
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/285,447
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/287,060
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/288,543
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 22
; LENGTH: 4208
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7671089CB1
; US-10-468-125-22

```

```

Query Match 99.3%; Score 851.6; DB 17; Length 4208;
Best Local Similarity 99.5%; Pred. No. 1.1e-234;
Matches 854; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGCGGCGCTCGGCACCCCTCCGCTTGGCGGGCCACCCACCGCGTGGTGGCGGG 60
Db 368 ATGGCGGCGCTCGGCACCCCTCCGCTTGGCGGGCCACCCACCGCGTGGTGGCGGG 427
QY 61 CTACCCGAGTGGCTCTGCGAGCAGCGCTGAGAGCGCCGAGGCGAGAGTGGCGTC 120
Db 428 CTACCCGAGTGGCTCTGCGAGCAGCGCTGAGAGCGCCGAGGCGAGAGTGGCGTC 487
QY 121 GCCCGCGGAAACCGCAGCAGCAGCTCTAGCTGGGCGCTGCTGGGCGAGCAGCTGGGCGCTG 180
Db 488 GCCCGCGGAAACCGCAGCAGCAGCTCTAGCTGGGCGCTGCTGGGCGAGCAGCTGGGCGCTG 547
QY 181 CAGGTGGTGGAGCTCGCGGCGAGAGAGCTTCGCGACTGCTCTTCTGAGGAGCAGT 240
Db 548 CAGGTGGTGGAGCTCGCGGCGAGAGAGCTTCGCGACTGCTCTTCTGAGGAGCAGT 607
QY 241 GCCGTGTGTGCGAGGAGAGCGGCGCTCATCACCGACCCGCGGCGCGAGCGGAGGAG 300
Db 608 GCCGTGTGTGCGAGGAGAGCGGCGCTCATCACCGACCCGCGGCGCGAGCGGAGGAG 667
QY 301 GAGGTGACATGATGAAGAGCATTAGAAAATTCAGCTCAATATAGTAGAGATGAAA 360
Db 668 GAGGTGACATGATGAAGAGCATTAGAAAATTCAGCTCAATATAGTAGAGATGAAA 727
QY 361 GATGAAATGCAACTTTAGATGGGCGAGATGTTTATTCACAGCGAGAGATTTTGTG 420

```

Db 728 GATGAAATGCAACTTTAGATCGGAGATGTTTATTACAGGACAGAGAAATTTTGTG 787
Qy 421 GGCCTTTCCAAAGGACAAATCAACAGAGTGTCTGAAATCTTGGCTGATACATTTTAAGGAC 480
Db 788 GGCCTTTCCAAAGGACAAATCAACAGAGTGTCTGAAATCTTGGCTGATACATTTTAAGGAC 847
Qy 481 TATGAGTCTCCACAGTGCAGTGGCAGATGGTGGTTCATTTGAAGAGTTTCTGAGCATG 540
Db 848 TATGAGTCTCCACAGTGCAGTGGCAGATGGTGGTTCATTTGAAGAGTTTCTGAGCATG 907
Qy 541 GCTGGGCTTAACCTGATCGCAATGGGCTAGTGAATCTGCACAGAGGCCCTTAAGATC 600
Db 908 GCTGGGCTTAACCTGATCGCAATGGGCTAGTGAATCTGCACAGAGGCCCTTAAGATC 967
Qy 601 ATGCAACAGATGAGTACCAACCGCTACGACAACTCACTGTGCTGATGACATAGCAGCA 660
Db 968 ATGCAACAGATGAGTACCAACCGCTACGACAACTCACTGTGCTGATGACATAGCAGCA 1027
Qy 661 AACTGTATATATCTAAATATCCCAACAAAGGCGCTCTGCTCACCAGACCCCGGAA 720
Db 1028 AACTGTATATATCTAAATATCCCAACAAAGGCGCTCTGCTCACCAGACCCCGGAA 1087
Qy 721 GAGTATCCAGAAAGTCAAAAGGTTTATGAGAACTGAAGGACCATATGCTGATCCCGTG 780
Db 1088 GAGTATCCAGAAAGTCAAAAGGTTTATGAGAACTGAAGGACCATATGCTGATCCCGTG 1147
Qy 781 AGCATGTCTGAACTGGAAGTGGATGGCTGCTCACTGCTGCTCAGTTTAAATTAAC 840
Db 1148 AGCATGTCTGAACTGGAAGTGGATGGCTGCTCACTGCTGCTCAGTTTAAATTAAC 1207
Qy 841 AAGAAGGTAGACTCTCTGA 858
Db 1208 AAGAAGGTAGACTCTCTGA 1225

RESULT 3

US-10-205-219-83
; Sequence 83, Application US/10205219
; Publication No. US20030138803A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pincock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018200
; CURRENT APPLICATION NUMBER: US/10/205,219
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 83
; LENGTH: 3008
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: N-G, N-G dimethylarginine dimethylaminohydrolase
; NAME/KEY: misc feature
; LOCATION: 2411, 2412
; OTHER INFORMATION: n is a or g or c or t
US-10-205-219-83

Query Match 83.4%; Score 715.6; DB 15; Length 3008;
Best Local Similarity 89.6%; Pred. No. 1.6e-195;
Matches 769; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
Qy 1 ATGGCCCGGCTCGGCCACCCCTCGCTTCGGCGGGCCACCCACCGCTGGTGGCGGG 60
Db 432 ATGGCCCGGCTCGGCCACCCCTCGCTTCGGCGGGCCACCCACCGCTGGTGGCGGGCT 491

Qy 61 CTACCCAGTCTGCTGCGAGCAGCGCTGAGAAAGCCCAAGGGCGAGAGGTGGACGTC 120
Db 492 CCGCCCGAGTCCCTGTGCGCCACCGCTGAGGCGCTCCAGGGCGAAGAGGTGGATTTC 551
Qy 121 GCCCGCGGAAAGCGGACAGCACAGCTCTACGTGGGCGTGTGGGCGACCAAGCTGGGCGTG 180
Db 552 GCTCGGCTGAGCGGACAGCACAGCTCTACGTGGGCGTGTGGGCGACCAAGCTGGGCGTG 611
Qy 181 CAGGTGTGAGTGTGCGGCGGACAGAGCCTTCCGAGCTGCGTCTTCGTGGAGGAGCTG 240
Db 612 CAGGTGTGAGTGTGCGGCGGACAGAGCCTTCCGAGCTGCGTCTTCGTGGAGGAGCTG 671
Qy 241 GCCGTGTGTGCGGAGGAGCGGCGCTCATCACCGACCCCGGGCGCGAGCGGAGGAAG 300
Db 672 GCCGTGTGTGCGGAGGAGCGGCGCTCATCACCGCGCCCGGGCGCGCTAGCGCGAGGAG 731
Qy 301 GAGGTTCACATGATGAAAGAGCAATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAA 360
Db 732 GAGGTTCACATGATGAAAGAGGCTTTGGAAAACTTCAGCTCAACATAGTAGAGATGAAA 791
Qy 361 GATGAAAAATGCACTTTAGATGGGAGAGTGTATTATTCAGGACAGAGATTTTGTG 420
Db 792 GATGAAAAATGCACTTTAGATGGGAGAGCTCTATTTCAGGACAGAGATTTTGTG 851
Qy 421 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACATTTAAGGAC 480
Db 852 GGCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAGATCTTGGCTGATACATTTCAAGGAC 911
Qy 481 TATGAGTCTCCACAGTGCAGTGGCGAGATGGTTCATTTGAAGAGTTCGTGAGCATG 540
Db 912 TACGAGTTCACAGTCCCGGCGGATTTTGTGCAATTTAAAGAGTTCGTGAGCATG 971
Qy 541 GCTGGGCTTAACCTGATCGCAATGGGCTTAGTGAATCTGCACAGAAAGGCCCTTAAGATC 600
Db 972 GCTGGGCTTAACCTGATCGCAATAGGCTCCAGTGAATCTGGCGAGAGGCCCTCAAGATC 1031
Qy 601 ATGCAACAGATGAGTACCCCGCTACGACAACTCACTGTGCTGCTGATGATAGCAGCA 660
Db 1032 ATGCAACAGATGAGTACCCCGCTATGACAAAGCTCACTGTACCGGACGATGGCGGCC 1091
Qy 661 AACTGTATATATCTAAATATCCCAACAAAGGCGACGCTTGTGTGACCCGACCCCGGAA 720
Db 1092 AACTGTATATATCTAAATATCCCAACAAAGGCGATGTCTGTGACCCGACCCCGGAA 1151
Qy 721 CAGTATCCAGAAAGTCAAAAGTTCATGAGAACTGAAGGACCATATGCTGATCCCGTG 780
Db 1152 GAGTATCCAGAAAGTCAAAAGTTCATGAGAAAGCTCAAGGACCATCTACTGATCCCGTG 1211
Qy 781 AGCATGTCTGAACTGGAAAGTGGATGGCTGCTCAGCTGCTCAGTTTAAATTAAC 840
Db 1212 AGCAATCTGAGATGGAAAGTGGATGGCTGCTCAGCTGCTCAGTTTAAATTAAC 1271
Qy 841 AAGAAGGTAGACTCTCTGA 858
Db 1272 AAGAAGGTAGACTCTCTGA 1289

RESULT 4

US-10-252-157-328
; Sequence 328, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program

SEQ ID NO 328
LENGTH: 3778
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030190640A1 198087.1
FEATURE:
NAME/KEY: unsure
LOCATION: 1083-1106
OTHER INFORMATION: a, t, c, g, or other
US-10-252-157-328

Query Match 82.2%; Score 705.4; DB 15; Length 3778;
Best Local Similarity 99.7%; Pred. No. 1.6e-192;
Matches 717; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 141 CCAGCTCTACGTGGCGTGTGGCAGCAACTGGGCTGCAGTGTGTGAGTGGCGGC 200
DB 1 CCAGCTCTACGTGGCGTGTGGCAGCAACTGGGCTGCAGTGTGTGAGTGGCGGC 60

QY 201 CGACGAGAGCTTCGGGACTCGCTTCGTGGAGGAGTGGCGCTGTGGCGAGGAGC 260
DB 61 CGACGAGAGCTTCGGGACTCGCTTCGTGGAGGAGTGGCGCTGTGGCGAGGAGC 120

QY 261 GGCCCTCATCCGACCCGGGGCGCGAGCCGGAGGAGGTTGACATGATGAAGA 320
DB 121 GGCCCTCATCCGACCCGGGGCGCGAGCCGGAGGAGGTTGACATGATGAAGA 180

QY 321 AGCATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAGATGAATGCACTTTAGA 380
DB 181 AGCATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAGATGAATGCACTTTAGA 240

QY 381 TGGCGGAGATGTTTATTCACAGGAGAGATTTTTGTGGGCTTTCCAAAAGGACAAA 440
DB 241 TGGCGGAGATGTTTATTCACAGGAGAGATTTTTGTGGGCTTTCCAAAAGGACAAA 300

QY 441 T-CACAGAGGTGCTGAATCTTGGCTGATATCTTTAAGGATATGCACTTCCACAGTGC 499
DB 301 TCCAAAGAGGTGCTGAATCTTGGCTGATATCTTTAAGGATATGCACTTCCACAGTGC 360

QY 500 CAGTGGCAGATGGTGTGATTTGAAGAGTTCCTGAGCAGTGGGCTTAACCTGATCG 559
DB 361 CAGTGGCAGATGGTGTGATTTGAAGAGTTCCTGAGCAGTGGGCTTAACCTGATCG 420

QY 560 CAATTGGGTCTAGTGAATCTGCACAGAGGCGCTTAAGATCATGCAACAGATGAGTACC 619
DB 421 CAATTGGGTCTAGTGAATCTGCACAGAGGCGCTTAAGATCATGCAACAGATGAGTACC 480

QY 620 ACCGCTACGACAACTCACTGCTGCTGATGATGACAGCAAACTGTATATCTTAATA 679
DB 481 ACCGCTACGACAACTCACTGCTGCTGATGATGACAGCAAACTGTATATCTTAATA 540

QY 680 TCCCAACAAAGGCGACGCTTCTGCTGCACCGAACCCCGAAGAGTATCCAGAAAGTGCAA 739
DB 541 TCCCAACAAAGGCGACGCTTCTGCTGCACCGAACCCCGAAGAGTATCCAGAAAGTGCAA 600

QY 740 AGTTTATGAGAACTGAAGGACCATATGCTGATCCCGTGAAGCATGCTGAACCTGGAAA 799
DB 601 AGTTTATGAGAACTGAAGGACCATATGCTGATCCCGTGAAGCATGCTGAACCTGGAAA 660

QY 800 AGTGTGATGGGTGCTCACTCTGCTCAGTTTAAATTAACAAGAGGTAGACTCTCTGA 858
DB 661 AGTGTGATGGGTGCTCACTCTGCTCAGTTTAAATTAACAAGAGGTAGACTCTCTGA 719

RESULT 5

US-09-925-300-512
Sequence 512, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:
APPLICANT: Craig Ruben
APPLICANT: Steve Ruben

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA101
CURRENT APPLICATION NUMBER: US/09/925,300
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05988
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 512
LENGTH: 3651
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1283)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (3641)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (3650)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-512

Query Match 73.5%; Score 630.4; DB 9; Length 3651;
Best Local Similarity 99.7%; Pred. No. 6.7e-171;
Matches 642; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 215 CGGACTGCGCTTCGTGGAGGACGTGGCGGTGGTGGAGGAGCGGCCCTCATCACCC 274
DB 2 CGGACTGCGCTTCGTGGAGGACGTGGCGGTGGTGGAGGAGCGGCCCTCATCACCC 61

QY 275 GACCCGGGCGCGGAGCCGGAGGAGGAGTTGACATGATGAAGAAGCATTTAGAAAAAC 334
DB 62 GACCCGGGCGCGGAGCCGGAGGAGGAGTTGACATGATGAAGAAGCATTTAGAAAAAC 121

QY 335 TTCAGCTCAATATAGTAGAGATGAAGAATGAATAATGCACTTTAGATGGCGAGATGTT 394
DB 122 TTCAGCTCAATATAGTAGAGATGAAGAATGAATAATGCACTTTAGATGGCGAGATGTT 181

QY 395 TATTACAGCGAGAGAAATTTTGTGGGCTTTCCAAAAGGACAAATCAACAGAGTGTCTG 454
DB 182 TATTACAGCGAGAGAAATTTTGTGGGCTTTCCAAAAGGACAAATCAACAGAGTGTCTG 241

QY 455 AAATCTGCTGATACCTTTTAAGGACTATGCACTATGCACTGCGCAGTGGCAGATGGGT 514
DB 242 AAATCTGCTGATACCTTTTAAGGACTATGCACTATGCACTGCGCAGTGGCAGATGGGT 301

QY 515 TGCAATTTGAAGAGTTTCTGAGCATGGCTGAGCTAACTGATCGCAATTTGGTCTAGTG 574
DB 302 TGCAATTTGAAGAGTTTCTGAGCATGGCTGAGCTAACTGATCGCAATTTGGTCTAGTG 361

QY 575 AATCTGCACAGAGGCGCTTTAAGATCATGCAACAGATGATGACCACTGACGACAAAC 634
DB 362 AATCTGCACAGAGGCGCTTTAAGATCATGCAACAGATGATGACCACTGACGACAAAC 421

QY 635 TCACTGTGCTGATGATAGCAGCAAACTGTATATCTTAATATCCCCCAACAGAGGC 694
DB 422 TCACTGTGCTGATGATAGCAGCAAACTGTATATCTTAATATCCCCCAACAGAGGC 481

QY 695 ACCTCTTGTGTCACCGAACCCCGAAGAGTATCCAGAAAAGTGCAGAAAGTTTATGAGAAC 754
DB 482 ACCTCTTGTGTCACCGAACCCCGAAGAGTATCCAGAAAAGTGCAGAAAGTTTATGAGAAC 541

QY 755 TGAAGGACCATATGCTGATCCCGTGAAGCATGCTGAATGGAAGAGTGGATGGGCTGC 814
DB 542 TGAAGGACCATATGCTGATCCCGTGAAGCATGCTGAATGGAAGAGTGGATGGGCTGC 601

QY 815 TCACCTGTGCTCAGTTTAAATTAACAAGAGGTAGACTCTCTGA 858
DB 602 TCACCTGTGCTCAGTTTAAATTAACAAGAGGTAGACTCTCTGA 644


```
QY 449 GTGCT-GAATCTTGGCTGATACHTTTAAAGACTATGAGTCTCCACAGTCCAGTGGCA 507
Db 777 GCGTGGAGATCGTGGGGACACGTTCCGGACTTCGCCGCTCCACTGTGCCAGTCTCG 836
QY 508 GATGGTGTGATTTGAAGAGTTTCTGCAGCATGCTGGGCTTACCTGATCGCAATTGGG 567
Db 837 GGTCCCTCCCACTGGCGGTCTCTGGGCATGGGGGACCTCGCACTGTGTGGCAGGC 896
QY 568 TCTAGTGAATGTCAGAGAAGCCCTTAAGATCATGCAACAGATGATGACCCGCTAC 627
Db 897 AGCAGCGACGTCGCCAAAGGCTGTCCGGCAATGGCAGTGTGACAGATCAACCCATAT 956
QY 628 GACAACTCA-CTGTGCTGTATGATAGCAGCAAACTGTATATATCT 674
Db 957 GCCTCCCTGACCCCTCCAGATGACGAGCTGCTGACTGTCTTTCT 1004
```

```
RESULT 8
US-09-783-590-5309
; Sequence 5309, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5309
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (22)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (42)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (64)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (98)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (134)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (139)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (232)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (257)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (263)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (270)
; OTHER INFORMATION: n equals a,t,g, or c
```

```
; NAME/KEY: misc feature
; LOCATION: (272)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (312)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (319)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (326)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-590-5309
```

```
Query Match 18.2%; Score 156; DB 9; Length 332;
Best Local Similarity 84.3%; Pred. No. 1.6e-34;
Matches 253; Conservative 0; Mismatches 35; Indels 12; Gaps 7;
```

```
QY 383 GCGGAGATGTTTTATTCACAGCAGAGAAATTTTTGTGGGCTTTCCAAAAGGACAAATC 442
Db 7 GGAAGATGTTTTATNCAAGCAGCAGAGAAATTTTTMTGGGCTTTCCAAAAGGACAAATC 66
QY 443 AAGGAGGTCTGAAATCTTTGGCTGATACATTTAAAGGACTATGAGTCTCCACAGTGC-CCA 501
Db 67 AAGGAGGTCTGAAATCTTTGGCTGATACATTTNAAGACTATGAGTCTCCACAGTGC-CCA 126
QY 502 GTGGCAG-ATGGTGTGCATTTGAAGAGTTTCTG-CAGCATGGCTGGGCT--TAACTTGA 556
Db 127 GTGGCAGNATGNTTGCATTTAAAGAGTTTCTGCCAGCATGGTTGGGCTTGAACCTGAA 186
QY 557 TGCATTTGGTCTAGTGA--TCTGCACAGAGCCCTTAGATCATGCAACAGATGAG 614
Db 187 TGCATTTGGTGTAGTGAATCTGCCAGAGGCCCTTAAAGATNATGCAACAAATGAG 246
QY 615 TGA--CCACCGCTACGACAACTCACTGTGCC--TGATGACATAGCAGCAAACTGTATAT 670
Db 247 TGAACCAACCGNTAAGNCAAACTNANTGTGCTTGTGATGACATAGCAGCAAACTGTATAT 306
```

RESULT 9

```
US-09-918-995-9117
; Sequence 9117, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9117
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(472)
; OTHER INFORMATION: n = A,T,C or G
; US-09-918-995-9117
```

```
Query Match 17.3%; Score 148.4; DB 10; Length 472;
Best Local Similarity 61.3%; Pred. No. 2.9e-32;
Matches 239; Conservative 0; Mismatches 151; Indels 0; Gaps 0;
```

```
QY 113 TGAAGCTCGCCGCGGGAACGCGAGCACCAGCTTACGTGGGCTGTGGGAGCAAGC 172
Db 82 TGGATCTGGCCAAAGCTCAAGGAGACCGGGGGCTGGGAGGTAAACTGAGGCAACGAC 141
QY 173 TGGGGCTGCGAGGTGTGGAGCTCGCCGCGCAGAGAGCCTTCGGGACTCGGCTTCGTGG 232
```

Db 142 TGGGCTACAGTGTAGAACTGCCACCTGAGAGGATTCGCGTGGGACCGCTGCTT 201
Qy 233 AGGACGTGCGCGTGTGTCGAGGAGACGGCCCTCATCACCGACCGCGGCGCGAGCC 292
Db 202 GCGACACGGCCGAGATCCAAAGGGGACACGGCCCTTAATCAGCGGCGCTGGAGCCCGCTC 261
Qy 293 CGAGGAGGAGCTTACATGATGAAGAGCATTAGAAAACCTCAGCTCAATATAGTAG 352
Db 262 GTAGCGAGAGTGTGATGAGTCCGCAAGCCCTGCAAGACCTGGGCTCCGAATGTGG 321
Qy 353 AGATGAAGATGAAATGCACTTTAGATGGCGGAGATGTTTTATTACAGGCGAGAGAT 412
Db 322 AATAGGACAGAGACGGGACGCTGGATGGCACTGAGCTTCTTCCACGGCGGAGT 381
Qy 413 TTTTGTGGGCTTTCCAAAAGAGCAATCAACAGAGGTGCTGAATCTTGGCTGATCTT 472
Db 382 TTTTGTGGGCTTTCCAAAAGAGCAATCAACAGAGGTGCTGAATCTTGGCTGATCTT 441
Qy 473 TTAAGGACTATGCACTCTCCACAGTGCCAG 502
Db 442 TCGGGACTTCGCGCTTACACTGTGCCAG 471

RESULT 10
US-10-450-826-8/c
; Sequence 8, Application US/10450826
; Publication No. US20040101818A1
; GENERAL INFORMATION:
; APPLICANT: JI, Darren
; APPLICANT: Axelrod, Douglas W.
; APPLICANT: Cook, Jonathan S.
; APPLICANT: Jaiswal, Neelam
; APPLICANT: Bistein, Richard
; APPLICANT: Houghton, Adam
; APPLICANT: Mertz, Lawrence
; TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiation
; FILE REFERENCE: 04921-5039-WO
; CURRENT APPLICATION NUMBER: US/10/450,826
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 60/255,882
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/285,591
; PRIOR FILING DATE: 2001-04-24
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 8
; LENGTH: 83946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. AL078459
US-10-450-826-8

Query Match 17.2%; Score 148; DB 17; Length 83946;
Best Local Similarity 100.0%; Pred. No. 5.5e-31;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 596 AGATCATGCAACAGATGAGTGACACCGCTGACCAAACTCACTGCTGATGACATAG 655
Db 18612 AGATCATGCAACAGATGAGTGACACCGCTGACCAAACTCACTGCTGATGACATAG 18553
Qy 656 CAGCAACTGTATATCTAAATATCCCAAGAGGACGCTTGTGTCACCGAACCC 715
Db 18552 CAGCAACTGTATATCTAAATATCCCAAGAGGACGCTTGTGTCACCGAACCC 18493
Qy 716 CGGAAGATATCCAGAAAGTGCAAGGT 743
Db 18492 CGGAAGATATCCAGAAAGTGCAAGGT 18465

RESULT 11
US-09-918-995-14871

; Sequence 14871, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14871
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(452)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-14871

Query Match 17.0%; Score 146; DB 10; Length 452;
Best Local Similarity 58.6%; Pred. No. 1.4e-31;
Matches 242; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
Qy 98 CCAAGGCGGAGGAGTGGACGTGCGCCGCGGGAACGGCAGCACCAGCTCTACGTGGGCG 157
Db 34 CACTGGCGAGCAAAANNNNAATNNAAGNNAAGGAGACACGGGNGCTGGAGGNA 93
Qy 158 TGTGGCGACCAAGCTGGGCTGCTGAGTGTGAGCTGCGCGCGCAGAGAGCTTCCGG 217
Db 94 AACTGTGGCAACGACTTGTGCTACAGCTGCTGAACTGCCACCTGAGGAGTCAATTGCCGC 153
Qy 218 ACTGCTGCTTGTGGAGGAGTGGCGTGTGTCGAGGAGAGCGCCCTCATCACCGAC 277
Db 154 TGGACCGCTGCTTGGCGACACGCGCGTGATCCAAAGGGACACGGCCCTTAATCACGGCG 213
Qy 278 CCGGGCGCGCGAGCGCGGAGGAAGAGGTTGCATGATGAAGAAGCATTAGAAAACTTC 337
Db 214 CCGGAGCGCGCTGCTGAGCCAGAGTTCGATGGAGTCCGCAAGCCCTGCAAGACCTGG 273
Qy 338 AGCTCAATATAGTAGATGAAGATGAATAATGCACTTTAGATGGCGGAGATGTTTAT 397
Db 274 GGTCCGAATTTGTGAAATAGGAGACGAGAACGCGCTGGATGGCAGCTGACGTTCTCT 333
Qy 398 TCACAGCAGAGAAATTTTGTGGCGCTTTCCAAAGGACAAATCAACGAGGTGCTGAAA 457
Db 334 TCACCGCGCGGAGTTTTCGTAGGCTCTCCAAATGACCAATCACCGAGGAGCTGAGA 393
Qy 458 TCTTGGCTGATCTTTTAAAGACTATGCACTTCCACAGTCCAGTGCCAGTGCCAGAT 510
Db 394 TCGTGGCGGACACGTTCCGGGACTTCGCGCTTCCACTGTGCCAGTCTCGSGT 446

RESULT 12
US-09-918-995-19254
; Sequence 19254, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19254
; LENGTH: 403

```
US-09-814-353-20553
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-918-995-19254

Query Match 15.6%; Score 134.2; DB 10; Length 403;
Best Local Similarity 59.6%; Pred. No. 3.4e-28;
Matches 226; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

Qy 121 GCCCGCGGAAACGACAGCAGCAGCTCTACGTGGCGCTGCTGGCGAGCAAGCTGGGCTG 180
Db 24 GCCCAAGCTCAAAGGAGACAGCGGGTGTGGGAGGTAACCTGAGGCAACGACTGGGCTA 83
Qy 181 CAGGTGGTGGAGCTCCCGCGCCACAGAGCCTTCGGGACTCGCTTCCTGCTGGAGAGGTG 240
Db 84 CAGCTGCTAGACTGCCACCTGAGAGTCAITTCGCGTGGGACCGCTGCTGGCGACAG 143
Qy 241 GCCGTGGTGGAGAGAGCGCCCTCATCACCGACCGCGGGGCGCCGAGCGGAGAGAG 300
Db 144 GCCGTGATCCAGAGGAGACAGCGACCTAATACGCGCGCCCTGGAGCCCGCTCGTAGGCCA 203
Qy 301 GAGGTGGATCATGATGAAGAACATTAGAAAACCTTCAGCTCAATATAGTAGAGATGAAA 360
Db 204 CAGGTGATGAGTCCGAAAGCCCTGCAAGACCTGGGGCTCCGAAATGTGGAAATAGGA 263
Qy 361 GATGAAATGCAACTTTAGATGGCGAGATGTTTATTCAGGCGAGAGATTTTGTG 420
Db 264 GACGAGAACCGACGCTGATGGCACTGACGTTCTGTTTCAACCGCGCGGAGATTTTCAGA 323
Qy 421 GCCTTTCARAGACAAATCAAGAGTGTCTGAAATCTTGGCTGATCTTTTAAGGAC 480
Db 324 GGCCTCTCTTAATGACCAATCACCGAGGAGCTGAGATCGTGGCGACACGTTTCGGGAC 383
Qy 481 TATGAGTCTCCACAGTGC 499
Db 384 TTCGAGTCTTCACTGTGC 402

RESULT 13
US-09-814-353-20553/c
; Sequence 20553, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20553
; LENGTH: 3189
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1_2_3189
; OTHER INFORMATION: n = A,T,C or G

US-09-814-353-20553
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-918-995-19254

Query Match 13.9%; Score 119.4; DB 10; Length 3189;
Best Local Similarity 99.2%; Pred. No. 1.8e-23;
Matches 120; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 738 AAAGTGTATGAGAACTCAAGGACCATATGTCATCCCGTGAGCATCTCTGAACTGGA 797
Db 3187 AAAGTGTATGAGAACTCAAGGACCATATGTCATCCCGTGAGCATCTCTGAACTGGA 3128
Qy 798 AAAGTGTATGAGTGGCTGCTCCTCCTCCTGCTCAGTTTATTAACAAGAGGTAGACTCTG 857
Db 3127 AAAGTGTATGAGTGGCTGCTCCTCCTCCTGCTCAGTTTATTAACAAGAGGTAGACTCTG 3068
Qy 858 A 858
Db 3067 A 3067

RESULT 14
US-09-736-457-404/c
; Sequence 404, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-736-457-404

Query Match 9.5%; Score 81.2; DB 9; Length 239;
Best Local Similarity 58.8%; Pred. No. 5.1e-13;
Matches 140; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy 185 TGCTGAGCTGCCGCGGACGAGAGCCTTCCGAGCTGCTTCGTGGAGGAGCTGGCGG 244
Db 238 TGCTAGAACTGCCCACTGAGAGTCAATTCGCGTGGACCGCTGCTGGGACACGCGG 179
Qy 245 TGCTGTGCGAGGAGAGCGGCCCTCATACCCGACCCCGGCGCGGAGCCGAGGAGAGG 304
Db 178 TGATCCAAAGGGGACACGGCCCTAATCACGCGGCTTGGAGCCCGCTCGTAGGCCAGAG 119
Qy 305 TTGACATGATGAAGAGCATTAAGAAAACCTCAGCTCAATATAGTAGAGATGAAGATG 364
Db 118 TCATGAGAGTCCGCAAGCCCTGCAAGACCTGGGGCTCCGAATTTGGAAATAGGAGAG 59
Qy 365 AAAATGCAACTTTAGATGGCGGAGATGTTTATTCACGCGAGAGATTTTGTGGG 422
Db 58 AGAACGCGACGCTGGATGGCACTGACGTCTCTTTCACCGCGCGGAGTTTTCGTAGG 1

RESULT 15
US-09-902-941-404/c
; Sequence 404, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
```

APPLICANT: Henderson, Robert A.
APPLICANT: Wang, Tongtong
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Johnson, Jeffrey C.
APPLICANT: Retter, Marc W.
APPLICANT: Marnerakis, Margarita
APPLICANT: Carter, Darrick
APPLICANT: Ranger, Gary R.
APPLICANT: Vedrick, Thomas S.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: McNabb, Andria
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C17
CURRENT APPLICATION NUMBER: US/09/902,941
NUMBER OF SEQ ID NOS: 2002
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 404
LENGTH: 239
TYPE: DNA
ORGANISM: Homo sapiens
US-09-902-941-404

Query Match 9.5%; Score 81.2; DB 9; Length 239;
Best Local Similarity 58.8%; Pred. No. 5.1e-13;
Matches 140; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY	185	TGCTGGAGCTGCCCGCGGACGAGACCTTCGGAGTGGCTCTTCGTGGAGGACCTGGCGG	244
Db	238	TGCTAGAACTGCCACCTGAGGAGTCAATGCCGCTGGGACCGCTGTGGCGACACGGCGG	179
QY	245	TGGTGTGCGAGGAGACGGCCCTCATCACCGGACCGCGGGCGCGGAGGAGGAGG	304
Db	178	TGATCCAGGGGACACGGCCCTAATCACGGGCGCTGGAGCCCGCTCGTAGCCAGAGG	119
QY	305	TTGACATGATGAAGAAGCATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAGATG	364
Db	118	TCGATGGAGTCCGCAAGGCCCTGCAAGACCTGGGGCTCCGAAATTGTGAAATAGGAGCG	59
QY	365	AAATGCACTTTAGATGGCGGAGATGTTTATTCACAGGACAGAAATTTTGTGGG	422
Db	58	AGAACGGAGCTGGATGGCACTGACGTTCTTTTACCAGCGGGAGTTTTTCGTAGG	1

Search completed: June 7, 2004, 21:24:56
Job time : 430 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 20:25:50 ; Search time 404 Seconds
(without alignments)

9022.163 Million cell updates/sec

Title: US-09-889-733B-1

4

Perfect score: 858

Sequence: 1 atggccgctcggccacc.....acaagaagtagactctcta 858

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq 29Jan04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2003as:*
- 8: Geneseqn2003bs:*
- 9: Geneseqn2003cs:*
- 10: Geneseqn2004s:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	858	100.0	858	3 AAA47655	Aaa47655 Dimethyla
2	772	90.0	858	3 ADE71465	Ade71465 Human Thr
3	772	90.0	1633	4 AAH02307	Aah02307 Human she
4	772	90.0	1633	6 ABQ88102	Abq88102 Human ost
5	772	90.0	3932	7 ABZ35939	Abz35939 Human sec
6	772	90.0	4208	6 AAL41494	Aal41494 Drug meta
7	721	84.0	858	9 ADE71463	Ade71463 Human Met
8	610	71.1	3651	3 AAF16077	Aaf16077 Human pro
9	585	68.2	3778	9 ADE53981	Ade53981 Human pro
10	148	17.2	83946	6 AQQ88101	Aq88101 Human ost
11	114	13.3	6373	5 ABA14642	Abal4642 Human ner
12	114	13.3	6381	5 ABA14641	Abal4641 Human ner
13	85	9.9	420	3 AAC10265	Aac10265 Human sec
14	80	7.0	60	6 ABN36282	Abn36282 Human spl
15	54	6.3	3008	9 ADB53014	Adb53014 Primary r
16	54	6.3	3008	9 ADB79843	Adb79843 Rat N-G, N
17	23	2.7	2523	7 ACC69598	Acc69598 Rat sweet
18	23	2.7	2771	3 AAZ50743	Aaz50743 Rat senso
19	22	2.6	862	6 ABQ47888	Abq47888 Oligonuc
20	22	2.6	862	6 ABQ47889	Abq47889 Oligonuc
21	22	2.6	898	6 ABQ41319	Abq41319 Oligonuc
22	22	2.6	898	6 ABQ41318	Abq41318 Oligonuc
23	21	2.4	21	3 AAA47661	Aaa47661 Expressed

C	24	21	2.4	21	3	AAA47663	AAA47663 Expressed
	25	21	2.4	21	3	AAA47664	AAA47664 Expressed
	26	20	2.3	444	6	ABN89945	Abn89945 Mouse clo
	27	20	2.3	21470	4	ABK42270	Abk42270 Genomic s
	28	20	2.3	21470	8	ADB60426	Adb60426 Connectiv
	29	20	2.3	110000	6	ABX08336	Abx08336 Human pho
C	30	19	2.2	314	2	AAQ60489	Aaq60489 Human bra
	31	19	2.2	368	4	ABK42398	Abk42398 Genomic s
C	32	19	2.2	368	4	AAK68647	Aak68647 Human imm
	33	19	2.2	368	8	ADB60554	Adb60554 Connectiv
C	34	19	2.2	573	4	AAL46258	Aal46258 Murine gr
	35	19	2.2	573	4	AAL46259	Aal46259 Murine gr
	36	19	2.2	718	4	AAL46254	Aal46254 Murine gr
	37	19	2.2	718	4	AAL46255	Aal46255 Murine gr
	38	19	2.2	758	2	AAT62819	Aat62819 Murine gr
	39	19	2.2	777	4	AAL46224	Aal46224 Murine gr
C	40	19	2.2	954	7	ACA21957	Aca21957 Prokaryot
	41	19	2.2	2481	6	ABT10749	Abt10749 Human bre
	42	19	2.2	4412	9	ADD24915	Add24915 DNA encod
	43	19	2.2	7084	7	ACC43645	Acc43645 Nucleotid
	44	19	2.2	7263	4	AAI58179	Aai58179 Human pol
	45	19	2.2	7263	8	ADB48145	Adb48145 Novel hum

ALIGNMENTS

RESULT 1

AAA47655
ID AAA47655 standard; cDNA; 858 BP.

AC AAA47655;

XX 08-NOV-2000 (first entry)

XX Dimethylarginine dimethylaminohydrolase (DDAH1) coding sequence.

XX Dimethylarginine dimethylaminohydrolase; DDAH; DDAH1; DDAH2;

XX arginine deaminase; hyperlipidemia; renal failure; hypertension; cancer;

XX restenosis; atherosclerosis; schizophrenia; multiple sclerosis; cancer;

XX ischemia reperfusion injury; septic shock; multi organ failure;

XX arthritis; skin disorders; inflammatory cardiac disease; migraine;

XX infection; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..858

FT FT /*tag= a

FT FT /product= "DDAH1"

XX WO200044888-A2.

XX 03-AUG-2000.

XX 26-JAN-2000; 2000WO-GB000226.

XX 26-JAN-1999; 99GB-00001705.

XX 04-JUN-1999; 99GB-00013066.

XX (UNLO) UNIV COLLEGE LONDON.

XX Vallance PJT, Leiper JM, Whitley GSJ, Charles IG;

XX WBI; 2000-543392/49.

XX P-PSDB; AAB01477.

XX Novel methylarginase polypeptides and polynucleotides, used to identify modulators of them, which are used in the treatment of e.g. cancer, hypertension, and bacterial infections.

XX Claim 1; Page 55-56; 68pp; English.

CC Nucleotides encoding methylarginase polypeptides, vectors comprising
CC these nucleotides and the polypeptides themselves can be used in
CC medicaments for the treatment of hyperlipidemia, renal failure,
CC hypertension, restenosis after angioplasty, atherosclerosis,
CC complications of heart failure, schizophrenia, multiple sclerosis or
CC cancer. Modulators of the enzyme can be used in medicaments for the
CC treatment of ischemia-reperfusion injury of the brain or heart, cancer,
CC lethal hypertension in severe inflammatory conditions such as septic
CC shock or multi-organ failure, or local and systemic inflammatory
CC disorders including arthritis, skin disorders, inflammatory cardiac
CC disease, migraine, or microbial or bacterial infection
XX
SQ Sequence 858 BP; 217 A; 214 C; 256 G; 171 T; 0 U; 0 Other;
Query Match 100.0%; Score 858; DB 3; Length 858;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 858; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGGCTCGCCACCCCTCCCTTCGGCGGGCCACCCACGCGGTGGTGGGGG 60
DB 1 ATGGCGGCTCGCGCCACCCCTCCCTTCGGCGGGCCACCCACGCGGTGGTGGGGG 60
QY 61 CTACCCGAGTGGCTCTGCGACAGCGCTGAGAGCGCCAGGCGGAGGAGTGGAGTC 120
DB 61 CTACCCGAGTGGCTCTGCGACAGCGCTGAGAGCGCCAGGCGGAGGAGTGGAGTC 120
QY 121 GCCCGCGCGAAGCGGACAGCTCTACCTGGCGGCTGTGGCGACAGAGCTGGGGCTG 180
DB 121 GCCCGCGCGAAGCGGACAGCTCTACCTGGCGGCTGTGGCGACAGAGCTGGGGCTG 180
QY 181 CAGGTGGTGGAGCTGCGCGCGAGAGAGCTTCGGGACTGCGTCTTCGTGGAGGAGCTG 240
DB 181 CAGGTGGTGGAGCTGCGCGCGAGAGAGCTTCGGGACTGCGTCTTCGTGGAGGAGCTG 240
QY 241 GCCGTGGTGGAGAGAGCGGCTCATACCCGAGCGCGGGCGCGAGCCGAGGAAG 300
DB 241 GCCGTGGTGGAGAGAGCGGCTCATACCCGAGCGCGGGCGCGAGCCGAGGAAG 300
QY 301 GAGGTGACATGATGAAAGAGCATTAGAAATCTTCAGCTCAATATAGTAGAGTGA 360
DB 301 GAGGTGACATGATGAAAGAGCATTAGAAATCTTCAGCTCAATATAGTAGAGTGA 360
QY 361 GATGAAATGCAATTTAGATGGCGAGATGTTTATTCACAGGACAGAAATTTTGTG 420
DB 361 GATGAAATGCAATTTAGATGGCGAGATGTTTATTCACAGGACAGAAATTTTGTG 420
QY 421 GGCCTTTCCAAAGGACAAATCAACGAGGCTGCAATCTGGCTGATCTTTTAGGAC 480
DB 421 GGCCTTTCCAAAGGACAAATCAACGAGGCTGCAATCTGGCTGATCTTTTAGGAC 480
QY 481 TATGAGTCTCCAGTGGCAGATGGGTGCAATTTGAGAGTTCCTGCGAGCATG 540
DB 481 TATGAGTCTCCAGTGGCAGATGGGTGCAATTTGAGAGTTCCTGCGAGCATG 540
QY 541 GCTGGGCTTAACCTGATCGCAATTTGGTCTAGTGAATCTGCAAGAGGCGCTTAAGATC 600
DB 541 GCTGGGCTTAACCTGATCGCAATTTGGTCTAGTGAATCTGCAAGAGGCGCTTAAGATC 600
QY 601 ATGCACAGATGAGTGACACCGCTACGACAAATCTCACTGTGCTGTATGATAGCAGCA 660
DB 601 ATGCACAGATGAGTGACACCGCTACGACAAATCTCACTGTGCTGTATGATAGCAGCA 660
QY 661 AACTGTATATATCAATATATCCCAAGAGGCAAGGCAAGTGTGTCACGACCCCGGAA 720
DB 661 AACTGTATATATCAATATATCCCAAGAGGCAAGGCAAGTGTGTCACGACCCCGGAA 720
QY 721 GAGTATCAGAAAGTGCAAGGTTTATGAGAACTGAAGGACCATATGCTGATCCCGTG 780
DB 721 GAGTATCAGAAAGTGCAAGGTTTATGAGAACTGAAGGACCATATGCTGATCCCGTG 780
QY 781 AGCATGCTGAACGGAAGAGGTGATGGGCTGCTACCTGCTCAGTTTTAATTAAC 840
DB 781 AGCATGCTGAACGGAAGAGGTGATGGGCTGCTACCTGCTCAGTTTTAATTAAC 840

QY 841 AAGAAGGTAGACTCTCTGA 858
DB 841 AAGAAGGTAGACTCTCTGA 858

RESULT 2

ADE71465

ID ADE71465 standard; DNA; 858 BP.

XX ADE71465;

XX ADE71465;

XX ADE71465;

DT 29-JAN-2004 (first entry)

XX

DE Human Thr87 wild-type dimethylarginine dimethylaminohydrolase-1 DNA.

XX

KW Human: dimethylarginine dimethylaminohydrolase-1; DDAH 1; chromosome 1;

KW single nucleotide polymorphism; SNP; ADMA;

KW asymmetrical dimethyl-arginine; nitric oxide; cardiovascular disease;

KW coronary heart disease; cerebrovascular disease; hypertension; diabetes;

KW susceptibility; genotyping; therapeutic targeting; antidiabetic;

KW cardiant; cerebroprotective; hypotensive; gene therapy; gene; ds.

XX

OS Homo sapiens.

XX

PH Location/Qualifiers

FT 1..858

FT /tag= a

FT /product= "Wild-type (Thr87) DDAH 1"

FT replace(260,T)

FT /tag= b

FT /standard name= "Single nucleotide polymorphism (SNP)"

FT /note= "Results in Thr (C variant) or Met (T variant) at

FT position 87 of the protein"

XX

PN WO2003089638-A1.

XX

PD 30-OCT-2003.

XX

PP 11-APR-2003; 2003WO-FI000274.

XX

PR 19-APR-2002; 2002US-00125456.

XX

PA (JURI-) JURILAB LTD OY.

XX

PI Valkonen V, Salonen JT, Pirskanen M, Tuomainen T, Laakso J;

PI Laaksonen R;

XX

DR WPI; 2003-854121/79.

XX

DR P-PSDB; ADE71466.

XX

PT New nucleic acid encoding a variant dimethylarginine
PT dimethylaminohydrolase-1 (DDAH-1) protein, useful for treating diabetes,
PT and its vascular complications, e.g. coronary or cerebrovascular disease
PT or hypertension.

XX

PS Disclosure; SEQ ID NO 3; 37pp; English.

XX

The invention relates to a variant dimethylarginine
dimethylaminohydrolase-1 (DDAH 1; ADE71464) and a nucleic acid encoding
it (ADE71463). DDAHs regulate the metabolism of ADMA (asymmetrical
dimethyl-arginine), a naturally occurring nitric oxide synthase inhibitor
and nitric oxide, which are relevant in disorders such as cardiovascular
disease and diabetes. The DDAH 1 gene is located on chromosome 1, which
has been implicated in susceptibility to diseases such as familial
combined hyperlipidaemia, premature coronary artery disease (CAD), non
insulin-dependent (type 2) diabetes mellitus, and diastolic hypertension.
The variant DDAH 1 of the invention has Met at position 87, rather than
Thr, which is caused by a C to T change in exon 1 at a position
corresponding to base 250 in ADE71463 and ADE71465. Individuals hetero-
zygous for the Thr87Met allele have an increased risk of
developing cardiovascular diseases and diabetes. The invention also
relates to a DDAH 1 cDNA capture probe; a method and kit for diagnosing a

CC susceptibility to a cardiovascular disease and diabetes in an individual
CC by determining the DDAH 1 genotype; a method of treating diabetes or
CC vascular complications of diabetes by enhancing nitric oxide
CC availability, production or concentration; a method of targeting
CC treatment of cardiovascular disease and diabetes in a hypertensive
CC patient by determining the DDAH 1 genotype and treating them with a drug
CC that affects nitric oxide availability, production or metabolism; and a
CC transgenic animal which carries a human variant DDAH 1 nucleic acid
CC sequence. The nucleic acid molecules and polypeptides are useful for
CC treating cardiovascular disease (e.g., coronary heart disease,
CC cerebrovascular disease, and hypertension), and diabetes and its vascular
CC complications. The methods are useful for determining whether a patient
CC will benefit from treatment with a drug which affects nitric oxide
CC availability, production or metabolism; a drug which reduces ADMA
CC availability or concentration; or an agent which elevates DDAH
CC availability or concentration (such as DDAH agonist). The methods are
CC also useful for determining whether a patient will be at risk of adverse
CC effects if DDAH antagonists are administered. The present sequence
CC represents a nucleic acid encoding the "wild-type" (Thr87) DDAH 1.
XX
SQ Sequence 858 BP; 218 A; 213 C; 258 G; 169 T; 0 U; 0 Other;

Query Match 90.0%; Score 772; DB 9; Length 858;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 822; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	23	CCGCTTCGGCCGGCCGACCCACCGCGTGTGGGGCGGTACCCGAGTGGCTCTGCCAGC	82
Db	23	CCGCTTCGGCCGGCCGACCCACCGCGTGTGGGGCGGTACCCGAGTGGCTCTGCCAGC	82
Qy	83	ACGGCTGAGAGGCGCAAGGGGAGGAGTGAAGTCCCGCGCGGCAACGGCAGCAC	142
Db	83	ACGGCTGAGAGGCGCAAGGGGAGGAGTGAAGTCCCGCGCGGCAACGGCAGCAC	142
Qy	143	AGCTCTACGTCGGCGTGTGGGCGAGCAAGTGGGGCTGCGAGTGGTGGAGCTGCCGGCG	202
Db	143	AGCTCTACGTCGGCGTGTGGGCGAGCAAGTGGGGCTGCGAGTGGTGGAGCTGCCGGCG	202
Qy	203	ACGAGAGCCTTCGCGACTTCTGCTGAGAGACGTCGCGCGTGGTGGAGGAGAGCG	262
Db	203	ACGAGAGCCTTCGCGACTTCTGCTGAGAGACGTCGCGCGTGGTGGAGGAGAGCG	262
Qy	263	CCCTCATCACCGACCGCGGCGCGCGAGCGGAGGAGGTTGACATGATGAAGAAG	322
Db	263	CCCTCATCACCGACCGCGGCGCGCGAGCGGAGGAGGTTGACATGATGAAGAAG	322
Qy	323	CATTAGAAAACTTCAGCTCAATATATAGATGAAAGATGAAATGCAACTTTAGATG	382
Db	323	CATTAGAAAACTTCAGCTCAATATATAGATGAAAGATGAAATGCAACTTTAGATG	382
Qy	383	CGCGAGATGTTTATTCACAGGCGAGAGATTTTCTGGGCGCTTCCAAAGGACAAATC	442
Db	383	CGCGAGATGTTTATTCACAGGCGAGAGATTTTCTGGGCGCTTCCAAAGGACAAATC	442
Qy	443	AACGAGTGTGAAATCTTGGCTGATACCTTTAAGGACTATGAGTCTCCACAGTCCAG	502
Db	443	AACGAGTGTGAAATCTTGGCTGATACCTTTAAGGACTATGAGTCTCCACAGTCCAG	502
Qy	503	TGGCAGATGGGTTGCAATTTGAAAGTTTCTGAGCATGGCTGGGCGCTAACCTGATCGAA	562
Db	503	TGGCAGATGGGTTGCAATTTGAAAGTTTCTGAGCATGGCTGGGCGCTAACCTGATCGAA	562
Qy	563	TTGGGCTCTAGTGAATCTGCACAGAGCCCTTAAGATCATGCAACAGATGAGTGACCC	622
Db	563	TTGGGCTCTAGTGAATCTGCACAGAGCCCTTAAGATCATGCAACAGATGAGTGACCC	622
Qy	623	GCTACGACAAACTCAGTGTGCTGATGACATACAGCAAACTGTATATATATATATCC	682
Db	623	GCTACGACAAACTCAGTGTGCTGATGACATACAGCAAACTGTATATATATATATCC	682
Qy	683	CCAAAGAGGCGAGCTTCTGCTGCAAGCCCGGAGAGTATCAGAAAGTGCAAGG	742
Db	683	CCAAAGAGGCGAGCTTCTGCTGCAAGCCCGGAGAGTATCAGAAAGTGCAAGG	742

Qy	743	TTTATGAGAACTCAAGGACCATATGTCCTCCCGTGGAGCATGCTGAACTGGAAGG	802
Db	743	TTTATGAGAACTCAAGGACCATATGTCCTCCCGTGGAGCATGCTGAACTGGAAGG	802
Qy	803	TGGATGGGCTCTCACTCTGCTGCTGCTCAAGTTTAAATTAACAAGAA	845
Db	803	TGGATGGGCTCTCACTCTGCTGCTGCTCAAGTTTAAATTAACAAGAA	845

RESULT 3

AAH02907
ID AAH02907 standard; DNA; 1633 BP.
XX
AC AAH02907;
DT
XX 15-JUN-2001 (first entry)
XX
DE Human shear stress-response coding sequence SEQ ID NO: 67.
XX
KW Human; shear stress-response protein; vascular disease; arteriosclerosis;
ds.
XX Homo sapiens.
XX
FN WO200125427-A1.
PD 12-APR-2001.
XX
PF 02-OCT-2000; 2000WO-JP006840.
PR 01-OCT-1999; 99JP-00280976.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
PA (NOJII) NOJIMA H.
PI Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;
PI Kuga T, Sekine S, Nakamura Y, Sugano S;
XX
DR WPI; 1-266308/27.
DR P-PSDB; AAB90784.
XX

DNA sequences, proteins encoded by them and antibodies against them
useful in diagnosis and treatment of vascular disease caused by
arteriosclerosis.
XX Claim 20; Page 404-407; 678pp; Japanese.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human shear stress response proteins. These are useful in the
CC diagnosis, treatment and screening of vascular diseases caused by
CC arteriosclerosis, including heart failure, post-PTCA restenosis and
CC hypertension
XX
SQ Sequence 1633 BP; 400 A; 406 C; 432 G; 393 T; 0 U; 2 Other;

Query Match 90.0%; Score 772; DB 4; Length 1633;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 822; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	23	CCGCTTCGGCCGGCCGACCCACCGCGTGTGGGGCGGTACCCGAGTGGCTCTGCCAGC	82
Db	345	CCGCTTCGGCCGGCCGACCCACCGCGTGTGGGGCGGTACCCGAGTGGCTCTGCCAGC	404
Qy	83	ACGCGCTGAGAGGCGCAAGGGGAGGAGTGGAGTGGCGCGCGGAAACGGCAGCAC	142
Db	405	ACGCGCTGAGAGGCGCAAGGGGAGGAGTGGAGTGGCGCGCGGAAACGGCAGCAC	464
Qy	143	AGCTCTACGTTGGGCGTGTGGGCGAGCAAGTGGGCGTGCAGGTGGTGGAGTGGCGGCG	202
Db	465	AGCTCTACGTTGGGCGTGTGGGCGAGCAAGTGGGCGTGCAGGTGGTGGAGTGGCGGCG	524
Qy	203	ACGAGAGCTTCCGAGCTGCGTCTTCGTGGAGGAGCTGGCGCTGGTGGTGGAGGAGCGG	262

Db 525 ACAGAGCCTTCGCGACTCGCTTCCTGAGGAGCTGCCGCTGCGAGGACGG 584
Qy 263 CCTCATCACCCGACCCGGGGCGCGAGCGAGGAGGTTGACATGATGAAGAAG 322
Db 585 CCTCATCACCCGACCCGGGGCGCGAGCGAGGAGGTTGACATGATGAAGAAG 644
Qy 323 CATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAATGCAACTTTAGATG 382
Db 645 CATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAATGCAACTTTAGATG 704
Qy 393 GGGAGATGTTTATTCACAGCAGAGAAATTTTGTGGGCTTCCAAAAGGACAAATC 442
Db 705 GGGAGATGTTTATTCACAGCAGAGAAATTTTGTGGGCTTCCAAAAGGACAAATC 764
Qy 443 AACGAGGTGCTCAATCTTGGCTGATCTTTAAGGACTATGCACTGCCACAGTGCCAG 502
Db 765 AACGAGGTGCTCAATCTTGGCTGATCTTTAAGGACTATGCACTGCCACAGTGCCAG 824
Qy 503 TGGCAGATGGTGTGATTTGAAGAGTTTCTGCAGCATGGCTGGGCTTAACCTGATCGCAA 562
Db 825 TGGCAGATGGTGTGATTTGAAGAGTTTCTGCAGCATGGCTGGGCTTAACCTGATCGCAA 884
Qy 563 TTGGGCTAGTGAATCTGCACAGAGGCGCTTAAGATCATGCAACAGATGATGACCAACC 622
Db 885 TTGGGCTAGTGAATCTGCACAGAGGCGCTTAAGATCATGCAACAGATGATGACCAACC 944
Qy 623 GTTACGACAACTCACTGTGCTGTGATGACATGACGACAACTGTATATATCTAATATCC 682
Db 945 GTTACGACAACTCACTGTGCTGTGATGACATGACGACAACTGTATATATCTAATATCC 1004
Qy 683 CCAACAAAGGGACGCTTGTGTCACCGAACCCCGAGAGATATCCAGAAAGTGAAGAAG 742
Db 1005 CCAACAAAGGGACGCTTGTGTCACCGAACCCCGAGAGATATCCAGAAAGTGAAGAAG 1064
Qy 743 TTTATGAGAACTGAAGGACCATATGCTGATCCCGTGAGCATGTCTGAACCTGGHAAAGG 802
Db 1065 TTTATGAGAACTGAAGGACCATATGCTGATCCCGTGAGCATGTCTGAACCTGGHAAAGG 1124
Qy 803 TGATGGGTGCTACCTGCTGCTGATGTTTAAATTAACAGAA 845
Db 1125 TGATGGGTGCTACCTGCTGCTGATGTTTAAATTAACAGAA 1167

RESULT 4
ABQ88102
ID ABQ88102 standard; cDNA; 1633 BP.
XX
AC ABQ88102;
DT 18-SEP-2002 (first entry)
XX
DE Human osteoblast differentiation related cDNA SEQ ID NO 9.
XX
XW Human; osteoblast; stem cell differentiation; bone tissue deposition;
KW osteoporosis; osteopathic; ss.
XX
OS Homo sapiens.
XX
FN WO20020301-A2.
XX
PD 27-JUN-2002.
XX
PF 18-DEC-2001; 2001WO-US048276.
XX
PR 18-DEC-2000; 2000US-0255882P.
PR 24-APR-2001; 2001US-0285691P.
XX
PA (GENE-) GENE LOGIC INC.
PA (PROC) PROCTER & GAMBLE CO.
XX
PI Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
Mertz L;

XX WPI; 2002-557663/59.
DR
XX
PT Use of genes and their expression profiles associated with osteoblast
PT differentiation for screening modulators bone formation, for diagnosing
PT or treating e.g. osteoporosis, or as markers for the differentiation
PT process.
XX
PS Claim 1; SEQ ID NO 9; 78pp + Sequence Listing; English.
PS
XX
CC The invention relates to genes and their expression profiles are used
CC for: (a) screening modulators of precursor stem cell differentiation into
CC osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
CC deposition of bone tissue, abnormal rate of osteoblast formation or
CC osteoporosis; or (c) treating or monitoring treatment of the conditions
CC cited in (b), or monitoring the progression of bone tissue deposition.
CC Specific conditions include postmenopausal osteoporosis, glucocorticoid
CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-
CC induced abnormalities in bone formation or bone loss, conditions that
CC involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
CC skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
CC or fibrous dysplasia. The present sequence is that of an osteoblast
CC differentiation associated cDNA marker of the invention. Note: The
CC sequence data for this patent did not form part of the invention. The
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1633 BP; 400 A; 406 C; 432 G; 393 T; 0 U; 2 Other;
Query Match 90.0%; Score 772; DB 6; Length 1633;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 822; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 23 CCGCTTTCGGCCGGGCGACCCACGCGTGTGCGGGCGGTACCCGAGTCGCTTCGCAGC 82
Db 345 CCGCTTTCGGCCGGGCGACCCACGCGTGTGCGGGCGGTACCCGAGTCGCTTCGCAGC 404
Qy 83 ACAGGCTGAGAGCGCCGAGGCGAGGAGTGGACCTCCCGCGCGGAGCGACGACACC 142
Db 405 ACAGGCTGAGAGCGCCGAGGCGAGGAGTGGACCTCCCGCGCGGAGCGACGACACC 464
Qy 143 AGCTCTACGTGGGCGTCTGCGGCGAGCAAGCTGGGGCTGAGGTGGAGCTGCCGGCG 202
Db 465 AGCTCTACGTGGGCGTCTGCGGCGAGCAAGCTGGGGCTGAGGTGGAGCTGCCGGCG 524
Qy 203 ACAGAGCCTTCGGACTCGTCTTCTGAGGAGACGTCGCGCTGCTGCGAGGACGG 262
Db 525 ACAGAGCCTTCGGACTCGTCTTCTGAGGAGACGTCGCGCTGCTGCGAGGACGG 584
Qy 263 CCTCATCACCCGACCCGGGGCGCGAGCGAGGAGGTTGACATGATGAAGAAG 322
Db 585 CCTCATCACCCGACCCGGGGCGCGAGCGAGGAGGTTGACATGATGAAGAAG 644
Qy 323 CATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAATGCAACTTTAGATG 382
Db 645 CATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAATGCAACTTTAGATG 704
Qy 383 GCGGAGATGTTTATTACAGGCGAGAGATTTTGTGGGCTTCCAAAAGGACAAATC 442
Db 705 GCGGAGATGTTTATTACAGGCGAGAGATTTTGTGGGCTTCCAAAAGGACAAATC 764
Qy 443 AACGAGGTGCTCAATCTTGGCTGATCTTTAAGGACTATGCACTGCCACAGTGCCAG 502
Db 765 AACGAGGTGCTCAATCTTGGCTGATCTTTAAGGACTATGCACTGCCACAGTGCCAG 824
Qy 503 TGGCAGATGGTGTGATTTGAAGAGTTTCTGCAGCATGGCTGGGCTTAACCTGATCGCAA 562
Db 825 TGGCAGATGGTGTGATTTGAAGAGTTTCTGCAGCATGGCTGGGCTTAACCTGATCGCAA 884
Qy 563 TTGGGCTAGTGAATCTGCACAGAGGCGCTTAAGATCATGCAACAGATGATGACCAACC 622
Db 885 TTGGGCTAGTGAATCTGCACAGAGGCGCTTAAGATCATGCAACAGATGATGACCAACC 944

naturally occurring polynucleotide sequence at least 90 % identical to the polynucleotide sequence, a polynucleotide complementary to them or an RNA equivalent of them. The polypeptide or polynucleotide are useful for treating, preventing or diagnosing a disease or condition associated with the expression of functional SRM. These are particularly useful for diagnosing, treating or preventing autoimmune/inflammatory disorders (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's disease), neurological disorders (e.g. epilepsy, Huntington's disease, dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease, multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety, schizophrenia or amnesia), or cell proliferative disorders (e.g. psoriasis, polycythemia vera, or cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain, breast, cervix or prostate). Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

Seq	Sequence	3932 BP;	1089 A;	893 C;	775 G;	1185 T;	0 U;	0 Other;	
Query Match	90.0%;	Score	772;	DB	7;	Length	3932;		
Best Local Similarity	99.9%;	Pred. No.	0;						
Matches	822;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	23	CGCCCTTCGGCGGGCCACCACGCGTGTGTGGGGCGCTACCGAGTGCCTCTGCCAGC	82						
DB	39	CGCCCTTCGGCGGGCCACCACGCGTGTGTGGGGCGCTACCGAGTGCCTCGGCCAGC	98						
QY	83	ACGCGCTGAGAGCGCCAGGGCGAGAGTGTGAAGCTGCGCGCGCGGCGGAACGGCAGCACC	142						
DB	99	ACGCGCTGAGAGCGCCAGGGCGAGAGTGTGAAGCTGCGCGCGCGGGAACGGCAGCACC	158						
QY	143	AGCTCTACGTGGCGGTGTGGGCAGCAAGCTGGGGTGCAGGTGGTGGAGCTGCCGGCGCG	202						
DB	159	AGCTCTACGTGGCGGTGTGGGCAGCAAGCTGGGGTGCAGGTGGTGGAGCTGCCGGCGCG	218						
QY	203	ACGAGAGCCTTCGGGACTGCGTCTTCGTGTGAGAGAGCTGGCGGTGGTGTGCGAGAGACGG	262						
DB	219	ACGAGAGCCTTCGGGACTGCGTCTTCGTGTGAGAGAGCTGGCGGTGGTGTGCGAGAGACGG	278						
QY	263	CCCTCATCACCCGACCCGCGGGCCCGAGCGCGAGGAAGAGGTTTGACATGATGAAGAAG	322						
DB	279	CCCTCATCACCCGACCCGCGGGCCCGAGCGCGAGGAAGAGGTTTGACATGATGAAGAAG	338						
QY	323	CATTAGAAAACCTTCAGCTCAATATAGTAGAGATGAAAGATGAAATGCAACTTTAGATG	382						
DB	339	CATTAGAAAACCTTCAGCTCAATATAGTAGAGATGAAAGATGAAATGCAACTTTAGATG	398						
QY	383	CGCGGAGATGTTTATTTACAGAGGACAGAGATTTTGTGGGCCCTTCCAAAGAGCAAAATC	442						
DB	399	CGCGGAGATGTTTATTTACAGAGGACAGAGATTTTGTGGGCCCTTCCAAAGAGCAAAATC	458						
QY	443	AACGAGGTGCTGAAATCTTGGCTGATACTTTTAAAGACTATGAGCTCCACAGTGCACG	502						
DB	459	AACGAGGTGCTGAAATCTTGGCTGATACTTTTAAAGACTATGAGCTCCACAGTGCACG	518						
QY	503	TGCGAGATGGGTTGCATTTCAAGAGTTTCTGACAGCATGGCTGGGCCCTTAACTGATCGCA	562						
DB	519	TGCGAGATGGGTTGCATTTCAAGAGTTTCTGACAGCATGGCTGGGCCCTTAACTGATCGCA	578						
QY	563	TTGGGTCTAGTGAATCTGCAAGAGGCCCTTTAAGATCATGCAACAGATGAGTGCACACC	622						
DB	579	TTGGGTCTAGTGAATCTGCAAGAGGCCCTTTAAGATCATGCAACAGATGAGTGCACACC	638						
QY	623	GCTACGCAAACTCAGTCGCTGATGACATAGCAGCAAACTGATATATCTAAATATCC	682						
DB	639	GCTACGCAAACTCAGTCGCTGATGACATAGCAGCAAACTGATATATCTAAATATCC	698						
QY	683	CCAAACAAAGGGCAGCTTTCGTGCAACGGAACCCCGGAAGAGTATCCGAAAAGTGCAAAGG	742						
DB	699	CCAAACAAAGGGCAGCTTTCGTGCAACGGAACCCCGGAAGAGTATCCGAAAAGTGCAAAGG	758						
QY	743	TTTATGAGAACTGAAAGGACCATATGCTGATCCCGGTGAGCATCTGAACTGGAAGG	802						

Db 759 TTTATGAGAACTGAAGGACCATATGCTGATCCCGTCAGCATGCTGAACTGAAAGG 818
Qy 803 TGGATGGCTGCTCACCTGCTGCTCAGTTTAAATTAACAGAA 845
Db 819 TGGATGGCTGCTCACCTGCTGCTCAGTTTAAATTAACAGAA 861

RESULT 6
ID AAL41494 standard; DNA; 4208 BP.
XX AAL41494;
AC AAL41494;
XX 21-NOV-2002 (first entry)
DT Drug metabolising enzyme encoding DNA - 7671089CBI.
DE
XX Anti-HIV; antiarteriosclerotic; dermatological; cytostatic; thyromimetic;
KW osteopathic; thrombolytic; ophthalmological; antilipemic; hepatotropic;
KW antidarrhoeic; antinflamatory; virucide; immunogenic; autoimmune;
KW drug metabolising enzyme; DME; inflammatory; AIDS; atherosclerosis;
KW contact dermatitis; cell-proliferative; cancer; cirrhosis; dwarfism;
KW developmental; hypochloridism; endocrine; osteoporosis; thrombosis;
KW diabetes; glaucoma; keratitis; metabolic; hyperlipidaemia; diarrhoea;
KW cystic fibrosis; gastrointestinal; gastroenteritis; liver; hepatitis;
KW Reye's syndrome; exogenous compound; gene therapy; enzyme; human; Gene;
ds.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 200..1225
CDS /*tag= a
/*product= "Drug metabolising enzyme protein"
XX
XX WO200266654-A2.
XX
XX 29-AUG-2002.
XX
XX 14-FEB-2002; 2002WO-US004918.
XX
XX 16-FEB-2001; 2001US-0269643P.
PR 23-FEB-2001; 2001US-0271332P.
PR 16-MAR-2001; 2001US-0276767P.
PR 06-APR-2001; 2001US-0282077P.
PR 19-APR-2001; 2001US-0285447P.
PR 27-APR-2001; 2001US-0287060P.
PR 03-MAY-2001; 2001US-0288543P.
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Astromoff A, Au-Young J, Baughn MR, Ding L, Duggan BM;
PI Forsythe JV, Gietzen KJ, Griffin JA, Lee EA, Lu Y, Richardson TW;
PI Ring HZ, Sanjanwala MM, Swarnakar A, Walla NK, Warren BA, Xu Y;
PI Yue H, Zebardjian Y;
XX
XX WPI; 2002-674949/72.
DR P-PSDB; AAC22798.
XX
XX New drug metabolizing enzymes (DME) useful for diagnosing, treating and
PT preventing diseases or conditions associated with aberrant DME
PT expression, e.g. cancer, AIDS, atherosclerosis, diabetes, glaucoma,
PT hepatitis, osteoporosis.
XX
XX Claim 5; Page 164-165; 166pp; English.
XX
XX The invention relates to an isolated polypeptide comprising 12 fully
CC defined sequences of 81-615 amino acids given in the specification; a
CC naturally occurring amino acid sequence at least 90% identical to, having
CC 81-599 amino acids, at least 93% identical to a sequence of 529 amino
CC acids, or at least 97% identical to a sequence of 615 amino acids, all
CC given in the specification; or a biologically active or immunogenic
CC fragment of the polypeptide. The polypeptides and polynucleotides are

CC useful in diagnosing, treating and preventing diseases or conditions
CC associated with the decreased expression or overexpression of a drug
CC metabolising enzyme (DME), such as autoimmune/inflammatory (e.g. AIDS,
CC atherosclerosis, contact dermatitis) cell-proliferative (e.g. cancer,
CC cirrhosis), developmental (e.g. dwarfism, hypothyroidism), endocrine
CC (e.g. osteoporosis, thrombosis, diabetes), eye (e.g. glaucoma,
CC keratitis), metabolic (e.g. hyperlipidaemia, cystic fibrosis),
CC gastrointestinal (e.g. gastroenteritis, diarrhoea), or liver (e.g.
CC hepatitis, Reye's syndrome) disorders. These are also useful in assessing
CC the effects of exogenous compounds on the expression of nucleic acid and
CC amino acid sequences of DME. The DME or its fragments are useful in
CC screening compounds for effectiveness as agonist or antagonist of the
CC polypeptides, or in altering the expression of the target polynucleotide
CC and compounds that specifically bind to or modulate the activity of the
CC polypeptide. The microarray is useful in monitoring or measuring protein-
CC protein interactions, drug-target interactions, and gene expression
CC profiles. The polynucleotides of the invention can be used in gene
CC therapy. This polynucleotide sequence represents the DNA encoding a drug
CC metabolising enzyme of the invention
XX
SQ Sequence 4208 BP; 1092 A; 1017 C; 883 G; 1216 T; 0 U; 0 Other;
Query Match 90.0%; Score 772; DB 6; Length 4208;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 822; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 23 CGCCCTTCGGCGCGCCACCCACGCGCTGTCGGCGGCGCTACCCGAGTCGCTCCCGAGC 82
Db 390 CGCCCTTCGGCGCGCGCCACCCACGCGCTGTCGGCGGCGCTACCCGAGTCGCTCCCGAGC 449
Qy 83 ACGCGCTGAGAAAGCCGCAAGGCGGAGGAGTGAGCTGCGCCGCGCGCAACGCGAGCACC 142
Db 450 ACGCGCTGAGAAAGCCGCAAGGCGGAGGAGTGAGCTGCGCCGCGCGCAACGCGAGCACC 509
Qy 143 AGCTCTACGTGGCGGTGTCGGCAGCAAGCTGGGCTGCGAGTGGAGTGGAGTGGCGGCG 202
Db 510 AGCTCTACGTGGCGGTGTCGGCAGCAAGCTGGGCTGCGAGTGGAGTGGAGTGGCGGCG 569
Qy 203 ACGAGAGCCTTCGGGACTGCGCTTCCTGTCGAGGAGAGCTGGCCCTGTCGCGAGGAGAGCG 262
Db 570 ACGAGAGCCTTCGGGACTGCGCTTCCTGTCGAGGAGAGCTGGCCCTGTCGCGAGGAGAGCG 629
Qy 263 CCTCTATCACCGACCCCGGGCGCGCGAGCGGAGGAGAGGTTGACATGATGAAGAAG 322
Db 630 CCTCTATCACCGACCCCGGGCGCGCGAGCGGAGGAGAGGTTGACATGATGAAGAAG 689
Qy 323 CATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAATGCAACTTTAGATG 382
Db 690 CATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAATGCAACTTTAGATG 749
Qy 383 GCGGAGATGTTTATTCACAGGACAGAGAAATTTTGTGGGCGCTTCCAAAAGGACAAATC 442
Db 750 GCGGAGATGTTTATTCACAGGACAGAGAAATTTTGTGGGCGCTTCCAAAAGGACAAATC 809
Qy 443 AACGAGGTGCTGAAATCTTGGGTGATACCTTTAAAGGACTATGACGTCTCCACAGTGCAC 502
Db 810 AACGAGGTGCTGAAATCTTGGGTGATACCTTTAAAGGACTATGACGTCTCCACAGTGCAC 869
Qy 503 TGGCAGATGGGTGCAATTTTGAAGAGTTTCTGAGCATGGCTGGGCGTAAACCTGATCGCAA 562
Db 870 TGGCAGATGGGTGCAATTTTGAAGAGTTTCTGAGCATGGCTGGGCGTAAACCTGATCGCAA 929
Qy 563 TTGGGTCTAGTGAATCTGACAGAGAGCCCTTAAAGATCATGCAACAGATGAGTGACACC 622
Db 930 TTGGGTCTAGTGAATCTGACAGAGAGCCCTTAAAGATCATGCAACAGATGAGTGACACC 989
Qy 623 GCTACGACAACTCAGTGTGCTGTATGATACATAGCAGCAAACTGTATATATATATATCC 682
Db 990 GCTACGACAACTCAGTGTGCTGTATGATACATAGCAGCAAACTGTATATATATATATCC 1049
Qy 683 CCACAAAGGCGACGCTTGTGTCACCGCAACCCCGGAGAGATTCAGAGAAAGTGCAAGG 742
Db 1050 CCACAAAGGCGACGCTTGTGTCACCGCAACCCCGGAGAGATTCAGAGAAAGTGCAAGG 1109

Db 623 GCTAGCAAACTCACTGTGCTGATGACATAGCAGCAAACTGTATATATCTAATATCC 682
Qy 683 CCAACAAAGGCGACGTCTTGTGTCACCGAACCCCGGAAGAGTATCCAGAAAGTGCAAGG 742
Db 683 CCAACAAAGGCGACGTCTTGTGTCACCGAACCCCGGAAGAGTATCCAGAAAGTGCAAGG 742
Qy 743 TTTATGAGAACTGAAGACCATATGTGATCCCGTGACATGCTGAACTGGAAGG 802
Db 743 TTTATGAGAACTGAAGACCATATGTGATCCCGTGACATGCTGAACTGGAAGG 802
Qy 803 TGGATGGGCTCTCACCTGTGCTCAGTCTTAAATTAACAAGAA 845
Db 803 TGGATGGGCTCTCACCTGTGCTCAGTCTTAAATTAACAAGAA 845

RESULT 8

AAFL16077
ID AAF16077 standard; cDNA; 3651 BP.
XX AC AAF16077;
XX 13-MAR-2001 (first entry)
XX Human prostate cancer antigen nucleotide sequence SEQ ID NO:512.
XX Human; prostate cancer; antigen; detection; diagnosis;
XX neuroprotective; cytosolic; cardiac; immunomodulatory; muscular;
XX vulnary; gastrointestinal; nephrotropic; anti-infective; gynaecological;
XX antibacterial; gene therapy; neural; immune; reproductive; renal;
XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX wound; infectious disease; ss.
XX Homo sapiens.
OS WO200055174-A1.
XX 21-SEP-2000.
XX 08-MAR-2000; 2000WO-US005988.
XX 12-MAR-1999; 99US-0124270P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX Rosen CA, Ruben SM;
XX WPI; 2000-587513/55.
XX P-PSDB; AAB56874.

Prostate cancer associated gene sequences, referred to as prostate cancer antigens, useful for treatment, prevention, and diagnosis of disorders such as prostate cancer.

Claim 1; Page 976-977; 2338pp; English.

AAFL15566 to AAF16505 encode the human prostate cancer associated proteins, called prostate cancer antigens, given in AAB56383 to AAB57302. The prostate cancer antigens can have neuroprotective, cytosolic, cardiac, immunomodulatory, muscular, vulnary, gastrointestinal, nephrotropic, anti-infective, gynaecological and antibacterial activities, and can be used in gene therapy. The prostate cancer antigen polynucleotides may be used for detection of prostate cancer, chromosome identification, as chromosome markers, and for numerous other diagnostic or research purposes. The prostate cancer antigens may be used to treat disorders such as neural, immune, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, and proliferative disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to AAB57303 represent sequences used in the exemplification of the present invention

Sequence 3651 BP; 1016 A; 801 C; 684 G; 1141 T; 0 U; 9 Other;

Query Match 71.1%; Score 610; DB 3; Length 3651;
Best Local Similarity 100.0%; Pred. No. 5.1e-304;
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 215 CGGACTGGGCTCTTCTGAGGACGCTGGCGCTGGTGTGGAGAGACGGCCCTCATCAGCC 274
DB 2 CGGACTGGGCTCTTCTGAGGACGCTGGCGCTGGTGTGGAGAGACGGCCCTCATCAGCC 61
QY 275 GACCGGGGGCGGAGCGGAG 334
DB 62 GACCGGGGGCGGAGCGGAG 121
QY 335 TTGAGCTCAATATAGTAGAGATGAAAGATGAAATGCAACTTTAGATGGCGAGAGATGTT 394
DB 122 TTGAGCTCAATATAGTAGAGATGAAAGATGAAATGCAACTTTAGATGGCGAGATGTT 181
QY 395 TATTCACAGGAGAGAAATTTTGTGGGCTTTTCAAAAGGACAAATCAACGAGGTGCTG 454
DB 182 TATTCACAGGAGAGAAATTTTGTGGGCTTTTCAAAAGGACAAATCAACGAGGTGCTG 241
QY 455 AAATCTTGGCTGATCTTTTAAAGGACTATGCAAGTCTCCACAGTGCAGTGGCAGATGGT 514
DB 242 AAATCTTGGCTGATCTTTTAAAGGACTATGCAAGTCTCCACAGTGCAGTGGCAGATGGT 301
QY 515 TGCATTTGAAGAGTTTCTGACAGCATGGCTGGGCTTAACCTGATCCCAATTTGGTCTAGTG 574
DB 302 TGCATTTGAAGAGTTTCTGACAGCATGGCTGGGCTTAACCTGATCGCAATTTGGTCTAGTG 361
QY 575 ATCTGTCACAGAGGCGCTTAAAGTATGACACAGATGAGTACACCGCTACGACAAAC 634
DB 362 AATCTGTCACAGAGGCGCTTAAAGTATGACACAGATGAGTACACCGCTACGACAAAC 421
QY 635 TCACCTGTGCTGATGACATAGCAGCAAACTGTATATATCTAAATATCCCAACAAAGGCG 694
DB 422 TCACCTGTGCTGATGACATAGCAGCAAACTGTATATATCTAAATATCCCAACAAAGGCG 481
QY 695 AGCTTGTGCTGACCGAACCGGAGAGATCCAGAAAGTCCAGAAAGTTCATGAGAAAC 754
DB 482 AGCTTGTGCTGACCGAACCGGAGAGATCCAGAAAGTTCATGAGAAAGTTCATGAGAAAC 541
QY 755 TGAAGGACCATATGCTGATCCCGTGAGCATGCTGAACTGGAAAAAGTGGATGGGCTGC 814
DB 542 TGAAGGACCATATGCTGATCCCGTGAGCATGCTGAACTGGAAAAAGTGGATGGGCTGC 601
QY 815 TCACCTGTGCTG 824
DB 602 TCACCTGTGCTG 611

RESULT 9

ADE53981
ID ADE53981 standard; cDNA; 3778 BP.
XX AC ADE53981;
XX 29-JAN-2004 (first entry)
XX Human prostate cancer cDNA #328.
XX Human; prostate cancer; ss; cDNA combination; differential expression; gene.
XX Homo sapiens.
XX US2003190640-A1.
XX 09-OCT-2003.
XX 29-MAY-2002; 2002US-00252157.
XX 31-MAY-2001; 2001US-0295048P.
XX (FARI/) FARIS M.
PA

PR	01-SEP-2000;	2000US-02232878P.
PR	01-SEP-2000;	2000US-02232943P.
PR	01-SEP-2000;	2000US-0223344P.
PR	01-SEP-2000;	2000US-02239345P.
PR	01-SEP-2000;	2000US-02239509P.
PR	05-SEP-2000;	2000US-02239513P.
PR	05-SEP-2000;	2000US-0230437P.
PR	06-SEP-2000;	2000US-0230438P.
PR	08-SEP-2000;	2000US-0231242P.
PR	08-SEP-2000;	2000US-0231243P.
PR	08-SEP-2000;	2000US-0231244P.
PR	08-SEP-2000;	2000US-0231414P.
PR	08-SEP-2000;	2000US-0231408P.
PR	08-SEP-2000;	2000US-0232081P.
PR	12-SEP-2000;	2000US-0231968P.
PR	14-SEP-2000;	2000US-0232397P.
PR	14-SEP-2000;	2000US-0233056P.
PR	21-SEP-2000;	2000US-0234223P.
PR	21-SEP-2000;	2000US-0234274P.
PR	25-SEP-2000;	2000US-0234997P.
PR	25-SEP-2000;	2000US-0234988P.
PR	25-SEP-2000;	2000US-0234848P.
PR	27-SEP-2000;	2000US-0233634P.
PR	27-SEP-2000;	2000US-0235836P.
PR	29-SEP-2000;	2000US-0236327P.
PR	29-SEP-2000;	2000US-0236367P.
PR	29-SEP-2000;	2000US-0236368P.
PR	29-SEP-2000;	2000US-0236369P.
PR	29-SEP-2000;	2000US-0236570P.
PR	02-OCT-2000;	2000US-0236802P.
PR	02-OCT-2000;	2000US-0237037P.
PR	02-OCT-2000;	2000US-0237038P.
PR	02-OCT-2000;	2000US-0237039P.
PR	02-OCT-2000;	2000US-0237040P.
PR	13-OCT-2000;	2000US-0239352P.
PR	13-OCT-2000;	2000US-0239937P.
PR	20-OCT-2000;	2000US-02405960P.
PR	20-OCT-2000;	2000US-0241785P.
PR	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241787P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	20-OCT-2000;	2000US-0246126P.
PR	20-OCT-2000;	2000US-02462221P.
PR	01-NOV-2000;	2000US-0244617P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0246523P.
PR	08-NOV-2000;	2000US-0246524P.
PR	08-NOV-2000;	2000US-0246525P.
PR	08-NOV-2000;	2000US-0246526P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246609P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0252100P.
PR	17-NOV-2000;	2000US-0252111P.

PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 05-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 03-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX useful for preventing, diagnosing and/or treating nervous system cancers
XX and metastases.
XX
XX Disclosure; SEQ ID NO 6973; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AB11004-AB11534) and proteins
XX (AB114678-AB118001) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemias; (d) wound healing
XX; (e) neurological diseases e.g. cerebral ischaemia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 6373 BP; 1810 A; 1396 C; 1180 G; 1987 T; 0 U; 0 Other;

Query Match 13.3%; Score 114; DB 5; Length 6373;
Best Local Similarity 100.0%; Pred. No. 4.3e-48;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 630 CAAACTCCTGCTGCTGATGACATAGCAGCAACTGTATATATCTAAATATCCCAACAA 689
DB 1 CAAACTCCTGCTGCTGATGACATAGCAGCAACTGTATATATCTAAATATCCCAACAA 60
QY 690 AGGCACGCTTCTGCTGACCGAACCCGGAAGTATCCAGAAAGTGCAGAGGT 743
DB 61 AGGCACGCTTCTGCTGACCGAACCCGGAAGTATCCAGAAAGTGCAGAGGT 114

RESULT 12
AB114641
ID AB114641 standard; DNA; 6381 BP.
XX
XX AC AB114641;
XX
XX DT 23-JAN-2002 (first entry)
XX
XX DE Human nervous system related polynucleotide SEQ ID NO 6972.
XX
XX KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antispasmodic; antianemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiatherogenic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
OS Homo sapiens.
XX
XX WO200159063-A2.
XX
XX PD 16-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US001334.
XX
XX PR 31-JAN-2000; 2000US-0179085P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217487P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236803P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0241827P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0244647P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.

PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 06-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
XX Disclosure; SEQ ID NO 6972; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AB11004-AB11534) and proteins
CC (AB14678-AB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 6381 BP; 1809 A; 1396 C; 1182 G; 1994 T; 0 U; 0 Other;
SQ
Query Match 13.3%; Score 114; DB 5; Length 6381;
Best Local Similarity 100.0%; Pred. No. 4.3e-48;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 630 CAAACTCACTGTGCTGATGACATAGCAGCAAACTGTATATATCTAAATATCCCAACAA 689
Db 1 CAAACTCACTGTGCTGATGACATAGCAGCAAACTGTATATATCTAAATATCCCAACAA 60
QY 690 AGGCACGTTTGTGTCACCGAACCCCGGAAGAGTATCCGAAAGTCCAAGTCAAGGT 743
Db 61 AGGCACGTTTGTGTCACCGAACCCCGGAAGAGTATCCGAAAGTCCAAGTCAAGGT 114
RESULT 13
AAC10265
ID AAC10265 standard; cDNA; 420 BP.
XX AAC10265;
XX
XX 06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 14340.
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

gene therapy; chromosome mapping; ss.
Homo sapiens.
EP1033401-A2.
06-SEP-2000.
21-FEB-2000; 2000EP-00200610.
26-FEB-1999; 99US-0122487P.
(GEST) GENSET.
Dumas Milne Edwards J, Duclert A, Giordano J;
WPI; 2000-500381/45.
New nucleic acid that is a 5' expressed sequence tag (5' EST) for
obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
diagnostic, forensic, gene therapy and chromosome mapping procedures.
Claim 1; SEQ ID NO 14340; 71pp + Sequence Listing; English.
The present sequence is one of a large number of 5' ESTs derived from
mRNAs encoding secreted proteins. No ORF has yet been conclusively
identified within the present sequence. The 5' ESTs were prepared from
total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
sequences usually correspond mainly to the 3' untranslated region (UTR)
of the mRNA because they are often obtained from oligo-dT primed cDNA
libraries. Such ESTs are not well suited for isolating cDNA sequences
derived from the 5' ends of mRNAs and even in those cases where longer
cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
ESTs are derived from mRNAs with intact 5' ends and can therefore be used
to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
diagnostic, forensic, gene therapy and chromosome mapping procedures.
They are used to obtain upstream regulatory sequences and to design
expression and secretion vectors
Sequence 420 BP; 131 A; 95 C; 94 G; 97 T; 0 U; 3 Other;
Query Match 9.9%; Score 85; DB 3; Length 420;
Best Local Similarity 100.0%; Pred. No. 3.7e-33;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 302 AGGTTGACATGATGAAGAAGCATTAGAAAACCTTCAGCTCAATATAGTAGAGATGAAG 361
Db 336 AGGTTGACATGATGAAGAAGCATTAGAAAACCTTCAGCTCAATATAGTAGAGATGAAG 395
Qy 362 ATGAAGATGCAACTTTAGATGCGG 386
Db 396 ATGAAGATGCAACTTTAGATGCGG 420
RESULT 14
ABN36282
ID ABN36282 standard; DNA; 60 BP.
XX
XX
AC ABN36282;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:9030.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
XX WO200210449-A2.
XX
PN
PD 07-FEB-2002.
XX

20-JUL-2001; 2001WO-IB001903.
XX
XX
PR 28-JUL-2000; 2000US-0221607P.
PR 02-MAY-2001; 2001US-0287724P.
XX
XX
PA (COMP-) COMPUGEN INC.
XX
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
XX
DR WPI; 2002-257383/30.
XX
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of a
PT genome, useful for detecting tissue-, pathology-, and developmental-
PT specific genes.
XX
XX
PS Example 1; SEQ ID NO 9030; 47pp; English.
XX
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-
CC)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises several
CC oligonucleotides, each capable of hybridising selectively to a set of
CC messenger RNAs transcribed from a given transcription unit of the genome,
CC which encodes one or more messenger RNA splice variants. The
CC oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a particular
CC biological or pathological state, and so allowing the detection of tissue
CC - and pathology-specific genes such as those genes only expressed in
CC specific tissue under a specific pathological condition; to detect
CC developmental specific genes; and to detect RNA transcripts and splice
CC variants of a transcriptome of a patient suffering from a particular
CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
CC rats, humans and mice, which are used in the exemplification of the
CC present invention. N.B. The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 60 BP; 18 A; 11 C; 19 G; 12 T; 0 U; 0 Other;
Query Match 7.0%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.8e-20;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 751 AAACCTGAAGGACCATATGCTGATCCCGTGAGCATGCTGAACTGGAAGGTGATGGG 810
Db 1 AAACCTGAAGGACCATATGCTGATCCCGTGAGCATGCTGAACTGGAAGGTGATGGG 60
RESULT 15
ADB53014
ID ADB53014 standard; DNA; 3008 BP.
XX
XX
AC ADB53014;
XX
XX
DT 04-DEC-2003 (first entry)
XX
XX
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3556.
XX
KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; ds.
XX
XX
OS Rattus norvegicus.
XX
XX
XX WO2003065993-A2.
XX
XX
PD 14-AUG-2003.
XX

PF 04-FEB-2003; 2003WO-US003482.
XX
XX 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378665P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
XX (GENE-) GENE LOGIC INC.
PA
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
XX Elashoff M;
PI
XX WFI; 2003-731472/59.
DR
XX
XX Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
XX Claim 44; SEQ ID NO 3556; 874pp; English.
PS
XX
XX The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX
SQ Sequence 3008 BP; 766 A; 842 C; 658 G; 740 T; 0 U; 2 Other;
Query Match 6.3%; Score 54; DB 9; Length 3008;
Best Local Similarity 100.0%; Pred No. 3.9e-17;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 136 CAGCACCAAGCTTACCTGGCGGTGCTGGGACAGCAAGCTGGCGCTCAGGTGTG 189
DB 567 CAGCACCAAGCTTACCTGGCGGTGCTGGGACAGCAAGCTGGCGCTCAGGTGTG 620

Search completed: June 7, 2004, 21:31:44
Job time : 407 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 20:34:35 ; Search time 2569 seconds
(without alignments)
9973.434 Million cell updates/sec

Title: US-09-889-733B-1
Perfect score: 858
Sequence: 1 atggcggcctggccaccc.....acaagaagtagactcctga 858

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_estc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_est3:*
- 12: gb_est4:*
- 13: gb_est5:*
- 14: gb_est6:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_pbg:*
- 27: em_gss_vrl:*
- 28: gb_gse1:*
- 29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	516	60.1	522	9	AI751463	AI751463 cn10b04.x
2	457	53.3	889	12	BI763990	BI763990 603049879
3	442	51.5	455	29	AX406329	AX406329 Homo sapi
4	350	40.8	356	12	BG977444	BG977444 RCS-C1016

5	340	39.6	455	29	AY406330	Pan trogl
6	323	37.6	624	10	AW514143	hg24802.x
7	269	31.4	535	9	AU279491	AU279491 AU279491
8	263	30.7	308	10	BF818963	BF818963 RCS-C1016
9	216	25.2	406	9	AI017468	AI017468 ou23c07.x
10	159	18.5	276	10	BF818967	RCS-C1016
11	138	16.1	1038	12	BG164988	602343844
12	115	13.4	312	9	AA376335	EST89005
13	106	12.4	700	28	AQ742098	HS 5566.B
14	59	6.9	167	10	BF933706	BF933706 CM2-WT017
15	59	6.9	357	13	BY233360	BY233360 BY233360
16	59	6.9	361	10	SB843865	SB843865 B8843865
17	59	6.9	387	13	BY292033	BY292033 BY292033
18	59	6.9	442	13	BY266349	BY266349
19	59	6.9	498	9	AI930353	ul63f11.y
20	59	6.9	507	9	AI225556	u105e03.y
21	59	6.9	539	9	AW012296	u07b05.y
22	59	6.9	550	9	AA986603	uc81h06.y
23	59	6.9	555	9	AI050531	AI050531 uc83d09.y
24	59	6.9	638	13	BY713422	BY713422 BY713422
25	59	6.9	685	14	CB247154	UI-M-P10-
26	59	6.9	702	13	BY731938	BY731938 BY731938
27	59	6.9	731	10	BB611533	BB611533 BB611533
28	59	6.9	759	12	BI736750	BI736750 603360448
29	59	6.9	777	14	CD349526	UI-M-PY0-
30	59	6.9	806	9	AU051351	AU051351 AU051351
31	59	6.9	844	9	AU051035	AU051035 AGENCOURT
32	59	6.9	862	13	BQ899468	BQ899468 AGENCOURT
33	59	6.9	923	12	BI332932	BI332932 602982027
34	59	6.9	1001	13	BY709864	BY709864 BY709864
35	59	6.9	1109	11	AQ010430	mus muscu
36	55	6.4	224	9	AA376525	EST888957
37	54	6.3	450	14	CB787749	AMGNNUC:N
38	54	6.3	493	14	CB725087	AMGNNUC:M
39	54	6.3	524	14	CB720352	AMGNNUC:N
40	48	5.6	547	9	AI661738	va31a06.x
41	48	5.6	874	12	BI560188	BI560188 603253311
42	47	5.5	338	13	BY331431	BY331431 BY331431
43	47	5.5	343	9	AA403666	va31a06.r
44	47	5.5	346	9	AA499371	VI89B12.r
45	47	5.5	388	13	BX490218	BX490218 DXF2P686H

ALIGNMENTS

RESULT 1
AI751463
LOCUS
DEFINITION
c10b04.x1 Normal Human Trabecular Bone Cells Homo sapiens CDNA
clone NHTBC_cn10b04 random, mRNA sequence.
ACCESSION
AI751463
VERSION
AI751463.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (bases 1 to 522)
Jia, L.B., Young, M.F., Touchman, J.W., Bouffard, G.G., Beckstrom-Sternberg, S.M., Green, E.D., Powell, J.I., Yang, L.M., Robey, P.G., Hotchkiss, R.N. and Francomano, C.A.
SGAP: The Skeletal Genome Anatomy Project
Unpublished (1997)
Contact: Libin Jia
Medical Genetics Branch
National Human Genome Research Institute
10/10C101, 9000 Rockville Pike, Bethesda, MP 20892-1267, USA
Tel: 301-402-4877
Fax: 301-496-7157
Email: libin@helix.nih.gov
DNA Sequencing and analyses by National Institutes of Health
Intramural Sequencing Center (NISC).

```
Plate: 10 row: b column: 04
Seq primer: -21M13 forward primer (ABI).
Location/Qualifiers
1..522
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NH78C-cn10b04"
/sex="Female"
/tissue_type="Bone"
/cell_type="Trabecular Bone Cells"
/lab_host="SURE"
/clone_lib="Normal Human Trabecular Bone Cells"
/notes="Organ: Hip; Vector: pBluescript; Site 1: EcoRI;
Library constructed by Dr. Marian Young and Dr. Pamela
Gehron Robey (NIDCR)"

FEATURES
source
ORIGIN
Query Match 60.1%; Score 516; DB 9; Length 522;
Best Local Similarity 100.0%; Pred. No. 1.9e-257;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 AGAAAACTTCAGCTCAATATAGTAGAGATGAAGATGAAATGCAACTTTAGATGGCGG 386
Db |||||
Qy 7 AGAAAACTTCAGCTCAATATAGTAGAGATGAAGATGAAATGCAACTTTAGATGGCGG 66
Db |||||
Qy 387 AGATGTTTTTATTCACAGGCAGAGAAATTTTGTGGGCTTTCCAAAAGGACAAATCAACG 446
Db |||||
Qy 67 AGATGTTTTTATTCACAGGCAGAGAAATTTTGTGGGCTTTCCAAAAGGACAAATCAACG 126
Db |||||
Qy 447 AGTGCTGAAATCTGGCTGATATTTTAAAGACTATGCACTCCAGTCCAGTCCAGTGGC 506
Db |||||
Qy 127 AGTGCTGAAATCTGGCTGATATTTTAAAGACTATGCACTCCAGTCCAGTCCAGTGGC 186
Db |||||
Qy 507 AGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGCTGGGCTACCTGATCGCAATTGG 566
Db |||||
Qy 187 AGATGGGTTGCATTTGAAGAGTTTCTGCAGCATGCTGGGCTACCTGATCGCAATTGG 246
Db |||||
Qy 567 GTCTAGTGAATCTGCACAGAGAGCCCTTAAAGATCATGCAACAGATGATGACCCGCTA 626
Db |||||
Qy 247 GTCTAGTGAATCTGCACAGAGAGCCCTTAAAGATCATGCAACAGATGATGACCCGCTA 306
Db |||||
Qy 627 GCACAAACTCAGTGTGCTGATGACATGACAGCAAACTGTATATCTAAATATCCCCAA 686
Db |||||
Qy 307 GCACAAACTCAGTGTGCTGATGACATGACAGCAAACTGTATATCTAAATATCCCCAA 366
Db |||||
Qy 687 CAAAGGGCAGCTTTGCTGCACCGAACCCCGGAGAGTATCCAGAAAGTCAAAAGTTTA 746
Db |||||
Qy 367 CAAAGGGCAGCTTTGCTGCACCGAACCCCGGAGAGTATCCAGAAAGTCAAAAGTTTA 426
Db |||||
Qy 747 TGAGAACTCAGGAGCATTATCTGATCCCGTGAGCATGTGTAACCTGGAAGAGTTGA 806
Db |||||
Qy 427 TGAGAACTCAGGAGCATTATCTGATCCCGTGAGCATGTGTAACCTGGAAGAGTTGA 486
Db |||||
Qy 807 TGGGCTGCTCAGCTGCTCAGTTTAAATTAACAA 842
Db |||||
Qy 487 TGGGCTGCTCAGCTGCTCAGTTTAAATTAACAA 522
Db |||||

RESULT 2
BI763990
LOCUS 603049879F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5189970 5',
DEFINITION mRNA sequence.
ACCESSION BI763990
VERSION BI763990.1 GI:15755568
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 889)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
```


Db 361 ATATGCTGATCCCGTGGAGCATGCTGACTGGAAGGATGGGCTGCTACCTGCT 420

QY 824 GCTCAGTTTAAATTAACAAGAA 845

Db 421 GCTCAGTTTAAATTAACAAGAA 442

RESULT 4

BG977444 356 bp mRNA linear EST 12-JUN-2001

LOCUS RCS-C10161-040101-013-D07 C10161 Homo sapiens cDNA, mRNA sequence.

DEFINITION BG977444

ACCESSION BG977444

VERSION BG977444.1 GI:14380179

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202653

PUBMED 10737800

COMMENT Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel.: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=RC5&t2=RC5-CI0161-040101-013-D07&t3=2001-01-04&t4=1>) Seq primer: puc 18 forward High quality sequence stop: 356.

FEATURES

Location/Qualifiers

1..356

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="C10161"

/note="Organ: colon; ins; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 40.8%; Score 350; DB 12; Length 356;

Best Local Similarity 100.0%; Pred. No. 6.8e-171;

Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 GAGGTGGAGTGGCCCGCGGACGCGAGCAGCAGCAGCTTACGTGGGCGTGGGCGAC 168

Db 7 GAGGTGGAGTGGCCCGCGGACGCGAGCAGCAGCAGCTTACGTGGGCGTGGGCGAC 66

QY 169 AAGCTGGGCGTGGAGTGGTGGAGTGGCGGCGGACGAGAGCTTCCGAGTGGCTTTC 228

Db 67 AAGCTGGGCGTGGAGTGGTGGAGTGGCGGCGGACGAGAGCTTCCGAGTGGCTTTC 126

Db 604 TGGCAGATGGGTTGCACTTTGAAGAGTTT 631

RESULT 3

AY406329 455 bp DNA linear GSS 15-DEC-2003

LOCUS Homo sapiens DDAH1 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

DEFINITION AY406329

ACCESSION AY406329.1 GI:39762303

VERSION AY406329

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 455)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES

Location/Qualifiers

1..455

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

<1..>455

/gene="DDAH1"

/locus_tag="HCM2521"

ORIGIN

Query Match 51.5%; Score 442; DB 29; Length 455;

Best Local Similarity 100.0%; Pred. No. 7.3e-219;

Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 404 GCAGAGATTTTGTGGGCTTTCCAAAAGGACAAATCAACAGGTGCTGAATCTTGG 463

Db 1 GCAGAGATTTTGTGGGCTTTCCAAAAGGACAAATCAACAGGTGCTGAATCTTGG 60

QY 464 CTGATCTTTTAAGGACTATGCACTCTCCACAGTGGCAGTGGGAGATGGTTGCATTGA 523

Db 61 CTGATCTTTTAAGGACTATGCACTCTCCACAGTGGCAGTGGGAGATGGTTGCATTGA 120

QY 524 AGATTCTTCGACATGGTGGGCTTAACCTGATCGCAATTTGGTCTAGTGAATCTGCAC 583

Db 121 AGATTCTTCGACATGGTGGGCTTAACCTGATCGCAATTTGGTCTAGTGAATCTGCAC 180

QY 584 AGAAGGCCCTTAAGATCATGCAACAGATGATGACCAACCGCTACGACAACTCACTGTGC 643

Db 181 AGAAGGCCCTTAAGATCATGCAACAGATGATGACCAACCGCTACGACAACTCACTGTGC 240

QY 644 CTGATGATAGCAGCAACTGTATATATCTAATATATCCCAACAAAGGCGACGTCTTGC 703

Db 241 CTGATGATAGCAGCAACTGTATATATCTAATATATCCCAACAAAGGCGACGTCTTGC 300

QY 704 TGCAACCAACCCCGAAGAGTATCCAGAAAGTGCAGAAAGTTTATGAGAACTCAAGGACC 763

Db 301 TGCAACCAACCCCGAAGAGTATCCAGAAAGTGCAGAAAGTTTATGAGAACTCAAGGACC 360

QY 764 ATATGCTGATCCCGTGGAGCATGTCTGAACTGGAAAAGGTGGATGGGCTGCTCACTGCT 823

Db 477 AAGAGTTTCTGAGCATGGCTGGGCTTACCTGATCGCAATGGTCTAGTGAATCTGCA 418
QY 583 CAGAGGCCCTTAAAGATCATGCAACAGATGAGTACCGCTACGACAACTCAGTGTG 642
Db 417 CAGAGGCCCTTAAAGATCATGCAACAGATGAGTACCGCTACGACAACTCAGTGTG 358
QY 643 CCTGATGACATAGCAGCAAACTGTATATATCTAAATATATCCCAAAAGGCGACGCTTTG 702
Db 357 CCTGATGACATAGCAGCAAACTGTATATATCTAAATATATCCCAAAAGGCGACGCTTTG 298
QY 703 CTGCACCGAACCCGGGAAGTATCCGAAAGTGCAGAGTGTATAGAACTGAAGGAC 762
Db 297 CTGCACCGAACCCGGGAAGTATCCGAAAGTGCAGAGTGTATAGAACTGAAGGAC 238
QY 763 CATATGCTGATCCCGTGGAGCATGCTGAACTGGAAGGTCAGGCTGCACCTGC 822
Db 237 CATATGCTGATCCCGTGGAGCATGCTGAACTGGAAGGTCAGGCTGCACCTGC 178
QY 823 TGCTCAGTTTTTAATTAACAAGAA 845
Db 177 TGCTCAGTTTTTAATTAACAAGAA 155

RESULT 7
A0279491
LOCUS
DEFINITION A0279491 CHONS1 Homo sapiens cDNA clone CHONS1000018 5', mRNA
sequence.
ACCESSION A0279491
VERSION A0279491.1 GI:28298718
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 535)
AUTHORS Inabayashi, H., Mori, T., Gojo, S., Kiyono, T., Sugiyama, T., Iris, R.,
Isogai, T., Hata, J., Tomoya, Y., and Umezawa, A.
TITLE Redifferentiation of dedifferentiated chondrocytes and
chondrogenesis of human bone marrow stromal cells via chondrosphere
formation with expression profiling by large-scale cDNA analysis
JOURNAL Exp. Cell Res. 288 (1), 35-50 (2003)
MEDLINE 22760598
PUBMED 12878157
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA Project, Sugiyama, T.; Wakamatsu, A.; Iris, R.;
Umezawa, A.; Fukuma, M.; Kusakari, S.; Hata, J.; Ishii, S.; Yamamoto, J.;
Isono, Y.; Saito, K.; Nakamura, Y.; Masuho, Y.; Nagai, K.; Isogai, T.
HRI human cDNA project; cDNA library construction & 5'-end one
pass sequencing; Helix Research Institute.
FEATURES
Location/Qualifiers
1..535
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CHONS1000018"
/cell_type="chondrocytes"
/clone_lib="CHONS1"
/note="Vector: pME18SFL3"

ORIGIN
Query Match 31.4%; Score 269; DB 9; Length 535;
Best Local Similarity 99.7%; Pred. No. 1.3e-128;
Matches 319; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 302 AGGTGACATGATGAAGAGCATTAGAAAACTTCAGTCAATATAGTAGAGATGAAG 361

Db 216 AGTTTGACATGATGAAGAGCATTAGAAAACTTCAGTCAATATAGTAGAGATGAAG 275
QY 362 ATGAAATGCAACTTTAGATGCGGAGATGTTTATTTCACAGGCGAGAGATTTTGTGG 421
Db 276 ATGAAATGCAACTTTAGATGCGGAGATGTTTATTTCACAGGCGAGAGATTTTGTGG 335
QY 422 GCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATATCTTTAAGGACT 481
Db 336 GCCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATATCTTTAAGGACT 395
QY 482 ATGCAAGTCTCCACAGTGCAGTGGCAGATGGTTCATCTTTGAAGAGTTCCTGCACATGG 541
Db 396 ATGCAAGTCTCCACAGTGCAGTGGCAGATGGTTCATCTTTGAAGAGTTCCTGCACATGG 455
QY 542 CTGGGCTTAACCTGATCCCAATTTGGTCTAGTGAATCTGCACAGAGCCCTTAAAGATCA 601
Db 456 CTGGGCTTAACCTGATCCCAATTTGGTCTAGTGAATCTGCACAGAGCCCTTAAAGATCA 515
QY 602 TGCAACAGATGAGTGACCAAC 621
Db 516 TGCAACAGATGAGTGACCAAC 535

RESULT 8
BF818963/c
LOCUS
DEFINITION RCS-C10160-221200-012-A08 C10160 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF818963
VERSION BF818963.1 GI:12156591
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 308)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC5&t2=RC5-C10160-
221200-012-A08&t3=2000-12-22&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 308.
Location/Qualifiers
1..308
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="C10160"

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="C10160"
/note="Organ: colon ins; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,715 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

```
ORIGIN
Query Match 30.7%; Score 263; DB 10; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 CAGGTGGTGGAGCTGCCGCGCCGACGAGAGCTTCCGACTGCGTCTTCGTCGAGGAGCGTG 240
DB 308 CAGGTGGTGGAGCTGCCGCGCCGACGAGAGCTTCCGACTGCGTCTTCGTCGAGGAGCGTG 249
QY 241 GCGTGGTGGAGGAGAGCGCGCTCATACCCGACCCGCGGCGCGCGAGCGAGGAG 300
DB 248 GCGTGGTGGAGGAGAGCGCGCTCATACCCGACCCGCGGCGCGCGAGCGAGGAG 189
QY 301 GAGGTTCATGATGAAGAAGCATTAGAAAACCTTCAGCTCAATATAGTAGAGATGAAA 360
DB 188 GAGGTTCATGATGAAGAAGCATTAGAAAACCTTCAGCTCAATATAGTAGAGATGAAA 129
QY 361 GATGAATGCAACTTAGATGGCGGAGATGTTTATTCACAGCGACAGAAATTTTGTG 420
DB 128 GATGAATGCAACTTAGATGGCGGAGATGTTTATTCACAGCGACAGAAATTTTGTG 69

QY 421 GCGCTTTCCAAAAGCACAATCA 443
DB 68 GCGCTTTCCAAAAGCACAATCA 46
```

```
RESULT 9
LOCUS AI017468 406 bp mRNA linear EST 27-AUG-1998
DEFINITION ou23c07.x1 Soares NPL T GBC S1 Homo sapiens cDNA clone
IMAGE:1627116.3, similar to TR:008557.008557
N-G-N-G-DIMETHYLARGININE DIMETHYLAMINOHYDROLASE. ; mRNA sequence.
AI017468
ACCESSION AI017468.1 GI:3231804
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 406)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 483 Std Error: 0.00
Seq primer: -40ml3 fwd: ET from Amersham.
```

```
Location/Qualifiers
1..406
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1627116"
/lab_host="DH10B"
/clone_lib="Soares NPL T GBC S1"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI CGAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
```

```
Soares and M. Fatima Bonaldo. "
Query Match 25.2%; Score 216; DB 9; Length 406;
Best Local Similarity 100.0%; Pred. No. 5.4e-101;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 630 CAAATCTACTGCTGCTGATGACATAGCAGCAAACTGTATATATCTAAATATATCCCCAACAA 689
DB 406 CAAATCTACTGCTGCTGATGACATAGCAGCAAACTGTATATATCTAAATATATCCCCAACAA 347
QY 690 AGGGCAGCTTCTGTCGACCGAACCCCGAGAGATCCAGAAAGTGAAGAGTTTATGA 749
DB 346 AGGGCAGCTTCTGTCGACCGAACCCCGAGAGATCCAGAAAGTGAAGAGTTTATGA 287
QY 750 GAAATCGAAGACCATATGCTGATCCCGTGAAGCATGTCTGAACCTGGAAGAGTGGATGG 809
DB 286 GAAATCGAAGACCATATGCTGATCCCGTGAAGCATGTCTGAACCTGGAAGAGTGGATGG 227
QY 810 GCTGCTCACTGCTGCTCAGTTTAAATTAACAGAA 845
DB 226 GCTGCTCACTGCTGCTCAGTTTAAATTAACAGAA 191
```

```
RESULT 10
LOCUS BF818967 276 bp mRNA linear EST 13-JAN-2001
DEFINITION RCS-C10160-221200-012-C11 C10160 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF818967
VERSION BF818967.1 GI:12156597
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 276)
AUTHORS Dias Neto,E.; Garcia Correa,R.; Verjovski-Almeida,S.; Briones,M.R.;
Nagai,M.A.; da Silva,W. Jr.; Zago,M.A.; Bordin,S.; Costa,F.F.;
Goldman,G.H.; Carvalho,A.F.; Matsukuma,A.; Baia,G.S.; Simpson,D.H.;
Brunstein,A.; deOliveira,P.S.; Bucher,P.; Jongeneel,C.V.;
O'Hare,M.J.; Soares,F.; Brentani,R.R.; Reis,L.F.; de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC5&t2=RC5-C10160-
221200-012-C11&t3=2000-12-22&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 270.
```

```
Location/Qualifiers
1..276
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="C10160"
/note="Organ: colon ins; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESIES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
```

[illegible]

QY		CTGCTCACCTGCTGTCA	828
Dδ		CTGCTCACCTGCTGTCA	138

RESULT 12	LOCUS	AA376335	312 bp	linear	EST 21-APR-1997
DEFINITION	EST99005 HSC172 cells II Homo sapiens cDNA 5' end, mRNA sequence.				
ACCESSION	AA376335				
VERSION	AA376335.1	GI:2028897			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo				

AUTHORS
Adams, W.D., Asilwagha, A.K., Fleischmann, N.D., Fulmer, R.A.,
Bult, C.J., Lee, N.H., Kirkness E.F., Weinstein, K.G., Gocayne, J.D.,
White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C.,
Clayton, R.A., Cline T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S.,
Glocke, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S. Jr.,

Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palanciano, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudke, D.M., Shirley, R., Small, K.V., Striggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarik, D., Cao, L., Cenada, M.A., Coleman, T.A., Collins, E.J.

Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,

Kozak, D. L., Kunsch, C., Hungjun, J., Li, H., Weissner, P. S., Olsen, H., Raymond, L., Wei, Y. F., Wing, J., Xu, C., Yu, G. L., Ruben, S. M.,

Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

JOURNAL Nature 377 (6547 Suppl.), 3-174 (1995)
MEDLINE 96036280
55555555
55555555

FUEMED
 /566098
 Other ESTs: THC99618
 Contact: Karlavag, AB
 COMMENT

Bioinformatics
The Institute for Genomic Research
Contact: Kelliavage, AN

THE INSTITUTE FOR GENOMIC RESEARCH
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056

Fax: 3018699423
Email: arkerlav@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.

```

FEATURES
Location/Qualifiers
1. .312
source

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
db_xref="MIMC (infect.181028"

```

```

/wd_xref="AIRC (almost):161023"
/db_xref="taxon:9606"
/cell_type="fibroblast"

```

```

/cell_line="HSC172 (60PDL)"
/dev_stage="fetal"

```

```
clone lib="HSC172 cells II"  
/notes=Organ: lung; Vector: pBluescript SK-; Site 1:
```

ECORI; Site_2: XhoI"

Query Match 13.4%; Score 115; DB 9; Length 312;

Best Local Similarity 100.0%; Pred. No. 2.5e-48;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 303 GGTGACATGATGAAGAAGCAATTGAAAACCTTCAGCTCAATATAGTAGAGATGAAGA 362

Db 161 GGTGACATGATGAAGAAGCAATTAGAAAACCTTCAGCTCAATATGATAGATGAAGA 220

QY 363 TGAATAATGCACTTTAGATGGCGAGATGTTTTATTCACAGCAGAGAAATTTTT 417

Db 221 TGAATAATGCACTTTAGATGGCGAGATGTTTTATTCACAGCAGAGAAATTTTT 275

RESULT 13

AQ742098/c

LOCUS

DEFINITION

HS 5566 B2 D01 SP6 RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1142 Col=2 Row=H, genomic survey sequence.

ACCESSION

AQ742098

VERSION

AQ742098.1

KEYWORDS

GSS.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 700)

MAHAIRES, G.G., WALLACE, J.C., SMITH, K., SWARTZELL, S., HOLZMAN, T., and KELLER, A., SHAKER, R., FURLONG, J., YOUNG, J., ZHAO, S., ADAMS, M.D., and HOOD, L.

AUTHORS

Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

TITLE

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

JOURNAL

99380599

MEDLINE

10449764

PUBMED

CONTACT: MAHAIRES GG, WALLACE JC, HOOD L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@redjones.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu

Plate: 1142 Row: H Column: 2

Seq primer: SP6

Class: BAC ends

High quality sequence stop: 700.

Location/Qualifiers

1..700

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/clone="Plate=1142 Col=2 Row=H"

/sex="male"

/clone_lib="RPCI-11 Human Male BAC Library"

/note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"

ORIGIN

Query Match 12.4%; Score 106; DB 28; Length 700;

Best Local Similarity 100.0%; Pred. No. 1.4e-43;

Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 740 AGGTTTATGAGAACTGAAGGACCATATGCTGATCCCGTGAGCATGCTGAACCTGAAA 799

Db 573 AGGTTTATGAGAACTGAAGGACCATATGCTGATCCCGTGAGCATGCTGAACCTGAAA 514

QY 800 AGGTGATGGTGTCTACCTCTCTCTGATTTTAAATTAACAGAA 845

Db 513 AGGTGATGGTGTCTACCTCTCTCTGATTTTAAATTAACAGAA 468

RESULT 14

BF933706

LOCUS

DEFINITION

CM2-NT0170-211200-647-e08 NT0170 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BF933706

VERSION

BF933706.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 167)

DIAS NETO, E., GARCIA CORREA, R., VARJOVSKI-ALMEIDA, S., BRIONES, M.R., NAGAI, M.A., DA SILVA, M. JR., ZAGO, M.A., BORDIN, S., COSTA, F.F., GOLDMAN, G.H., CARVALHO, A.F., MATSUKUMA, A., BAIA, G.S., SIMPSON, D.H., BRUNSTEIN, A., DEOLIVEIRA, P.S., BUCHER, P., JONGENSEL, C.V., O'HARE, M.J., SOARES, F., BRENTANI, R.R., REIS, L.F., DE SOUZA, S.J. and SIMPSON, A.J.

AUTHORS

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

TITLE

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

20202663

MEDLINE

10737800

PUBMED

CONTACT: SIMPSON A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-NT0170-211200-647-e08&t3=2000-12-21&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 27

High quality sequence stop: 166.

Location/Qualifiers

1..167

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="NT0170"

/notes="Organ: nervous tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 6.9%; Score 59; DB 10; Length 167;

Best Local Similarity 100.0%; Pred. No. 3.7e-19;

Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 739 AAGTTTATGAGAACTGAAGGACCATATGCTGATCCCGTGAGCATGCTGAACCTGGA 797

Db 55 AAGTTTATGAGAACTGAAGGACCATATGCTGATCCCGTGAGCATGCTGAACCTGGA 113

RESULT 15

BY233360

LOCUS

DEFINITION

BY233360 RIKEN full-length enriched, adult inner ear Mus musculus cDNA clone F930030D22 5', mRNA sequence.

ACCESSION

BY233360

VERSION

BY233360.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 357)

1 (bases 1 to 357)

ORIGIN

Query Match 6.9%: Score 59; DB 13; Length 357;

BEST LOCAL SIMILARITY 100.0%; FREQ. NO: 4.1E-13;
Matches 59: Conservative 0: Mismatches 0: Indels

1 ATGGCCGGCCTCGGCCACCCCTCCGCCCTCGGCCGGCCACCCACGCCGTGGTGGGC 59

Db
134 ATGGCCGGCCTCGGCCACCCCTCCGCCCTCGGCCGGGCCACCCACGCCGTGTCGGGC 192

Search completed: June 7, 2004, 23:11:51
Job time : 2572 secs

Analysis of the mouse transcriptome based on functional annotation

02-567778
Nure 420, 563-573 (2002)
Nature 420, 563-573 (2002)
22354683
12466851
Contact: Yoshihide Havashizaki

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsic.riken.go.jp
URL: <http://genome.gsc.riken.go.jp/>
Atzawa, K., Akimura, T., Arakawa, I., Carninci, P., Fukuda, S.,
Hiyazane, T., Imutani, K., Ishii, Y., Itoh, M., Kawaj, J., Konno, H.,
Miyazaki, A., Morita, M., Nakamura, M., Nomura, K., Numazaki, R.,
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
Shiraki, T., Tagami, M., Waki, K., Wataniki, A., Muramatsu, M. and
Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences *Mamm. Genome* 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. *Genome Res.* 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001). cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Kirk W. Beisel (Boys Town National Research Hospital 555 North 30th Street Omaha, NE 68131 USA) whose assistance we gratefully acknowledge. Please visit our web site (<http://genome.csc.riken.go.jp>) for further details.

Location/Qualifiers
1. .357

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 21:17:45 ; Search time 96 Seconds
(without alignments)
4959.874 Million cell updates/sec

Title: US-09-889-733B-1
Perfect score: 858
Sequence: 1 atggcggcctggccaccc.....acaagaagtgactcctga 858

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A COMB.seq*
2: /cgn2_6/prodata/2/ina/5B COMB.seq*
3: /cgn2_6/prodata/2/ina/6A COMB.seq*
4: /cgn2_6/prodata/2/ina/6B COMB.seq*
5: /cgn2_6/prodata/2/ina/PCFUS COMB.seq*
6: /cgn2_6/prodata/2/ina/backfiles1.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	2-2	7263	4	US-09-620-312D-55
2	18	2-1	1014	4	US-09-252-991A-9227
3	18	2-1	1588	3	US-08-617-785-7
4	18	2-1	1728	4	US-09-817-464-7
5	18	2-1	1728	4	US-09-134-000C-1052
6	18	2-1	1728	4	US-08-617-785-11
7	18	2-1	2745	3	US-08-617-464-11
8	18	2-1	2745	3	US-08-617-785-13
9	18	2-1	2765	3	US-08-617-464-13
10	18	2-1	2997	1	US-08-453-862-1
11	18	2-1	2997	2	US-08-452-734A-1
12	18	2-1	2997	3	US-08-176-401B-1
13	18	2-1	2997	5	PCT-US94-14989-1
14	18	2-1	3021	4	US-09-016-434-1118
15	18	2-1	3804	3	US-08-617-785-3
16	18	2-1	3804	4	US-08-617-464-3
17	18	2-1	1664976	4	US-08-916-421B-1
18	18	2-1	1830121	4	US-09-557-88A-1
19	18	2-1	1830121	4	US-09-643-990A-1
20	18	2-1	4403765	3	US-09-103-840A-2
21	18	2-1	4411529	3	US-09-103-840A-1
22	17	2-0	243	4	US-09-543-681A-3669
23	17	2-0	515	4	US-08-621-976-12843
24	17	2-0	524	4	US-08-621-976-1137
25	17	2-0	528	4	US-09-134-001C-1910
26	17	2-0	579	4	US-09-449-285A-13
27	17	2-0	586	3	US-09-328-111-588

28	17	2.0	696	4	US-09-252-991A-2735	Sequence 2735, Ap
29	17	2.0	792	4	US-09-252-991A-7444	Sequence 7444, Ap
30	17	2.0	807	4	US-09-252-991A-164	Sequence 164, App
31	17	2.0	891	4	US-09-252-991A-176	Sequence 176, App
32	17	2.0	948	4	US-09-252-991A-172	Sequence 172, App
33	17	2.0	960	4	US-09-252-991A-7256	Sequence 7256, Ap
34	17	2.0	1007	4	US-09-247-155-81	Sequence 81, Appli
35	17	2.0	1069	4	US-09-372-422A-7	Sequence 7, Appli
36	17	2.0	1088	4	US-09-148-545-132	Sequence 132, App
37	17	2.0	1143	4	US-09-148-545-79	Sequence 79, Appl
38	17	2.0	1168	4	US-09-484-970B-89	Sequence 89, Appl
39	17	2.0	1333	4	US-09-372-422A-9	Sequence 9, Appli
40	17	2.0	1354	4	US-09-620-312D-308	Sequence 308, App
41	17	2.0	1611	4	US-09-252-991A-183	Sequence 183, App
42	17	2.0	2338	3	US-09-232-200-66	Sequence 66, Appl
43	17	2.0	2338	4	US-09-232-197-66	Sequence 66, Appl
44	17	2.0	2338	4	US-09-232-201-66	Sequence 66, Appl
45	17	2.0	2338	4	US-09-232-195-66	Sequence 66, Appl

ALIGNMENTS

RESULT 1
US-09-620-312D-55
; Sequence 55, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chengnua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620 312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: DE_FL_genes Version 1.0
; SEQ ID NO 55
; LENGTH: 7263
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)...(5057)
US-09-620-312D-55

Query Match 2.2%; Score 19; DB 4; Length 7263;
Best Local Similarity 100.0%; Pred.No. 7.3;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 CTGCAGGTGGTGGAGCTGC 196

Db 1251 CTCGAGGTGGTGGAGCTGC 1269


```
RESULT 2
US-09-252-991A-9227
; Sequence 9227, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 9227
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9227

Query Match      2.1%; Score 18; DB 4; Length 1014;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 CTCGCGGCGGACGAGGC 210
DB 811 CTCGCGGCGGACGAGGC 828

RESULT 3
US-08-617-785-7
; Sequence 7, Application US/08617785E
; Patent No. 6228610
; GENERAL INFORMATION:
; APPLICANT: Flor, Peter J.
; APPLICANT: Kuhn, Ranier
; APPLICANT: Lindaur, Kristen
; APPLICANT: Puttner, Irene
; APPLICANT: Knopfel, Thomas
; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
; FILE REFERENCE: 4-19679/A/PCT
; CURRENT APPLICATION NUMBER: US/08/617,785E
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: PCT/EP94/02991
; EARLIER FILING DATE: 1994-09-07
; EARLIER APPLICATION NUMBER: EPO 9416553.7
; EARLIER FILING DATE: 1994-08-19
; EARLIER APPLICATION NUMBER: EPO 93810663.0
; EARLIER FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1588
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1447)
US-08-617-785-7

Query Match      2.1%; Score 18; DB 3; Length 1588;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 AACTTCAGCTCAATATAG 349
DB 228 AACTTCAGCTCAATATAG 245

RESULT 4
US-09-817-464-7
; Sequence 7, Application US/09817464
; Patent No. 6515107
; GENERAL INFORMATION:
; APPLICANT: Flor, Peter J.
; APPLICANT: Kuhn, Ranier
; APPLICANT: Lindaur, Kristen
; APPLICANT: Puttner, Irene
; APPLICANT: Knopfel, Thomas
; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
; FILE REFERENCE: 4-19679/A/PCT
; CURRENT APPLICATION NUMBER: US/09/817,464
; EARLIER FILING DATE: 2001-03-26
; EARLIER APPLICATION NUMBER: US/08/617,785
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: EPO 9416553.7
; EARLIER FILING DATE: 1994-08-19
; EARLIER APPLICATION NUMBER: EPO 93810663.0
; EARLIER FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1588
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1447)
US-09-817-464-7

Query Match      2.1%; Score 18; DB 4; Length 1588;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 AACTTCAGCTCAATATAG 349
DB 228 AACTTCAGCTCAATATAG 245

RESULT 5
US-09-134-000C-1052/C
; Sequence 1052, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; EARLIER FILING DATE: 1998-08-13
; EARLIER APPLICATION NUMBER: US 60/055,778
; EARLIER FILING DATE: 1997-08-15
; EARLIER APPLICATION NUMBER: EPO 93810663.0
; EARLIER FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1052
; LENGTH: 1728
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-1052

Query Match      2.1%; Score 18; DB 4; Length 1728;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 ATAGCAGCAAACTGTATA 669
DB 565 ATAGCAGCAAACTGTATA 548

RESULT 6
US-08-617-785-11
```

```
RESULT 4
US-09-817-464-7
; Sequence 7, Application US/09817464
; Patent No. 6515107
; GENERAL INFORMATION:
; APPLICANT: Flor, Peter J.
; APPLICANT: Kuhn, Ranier
; APPLICANT: Lindaur, Kristen
; APPLICANT: Puttner, Irene
; APPLICANT: Knopfel, Thomas
; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
; FILE REFERENCE: 4-19679/A/PCT
; CURRENT APPLICATION NUMBER: US/09/817,464
; EARLIER FILING DATE: 2001-03-26
; EARLIER APPLICATION NUMBER: US/08/617,785
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: EPO 9416553.7
; EARLIER FILING DATE: 1994-08-19
; EARLIER APPLICATION NUMBER: EPO 93810663.0
; EARLIER FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1588
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1447)
US-09-817-464-7

Query Match      2.1%; Score 18; DB 4; Length 1588;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 AACTTCAGCTCAATATAG 349
DB 228 AACTTCAGCTCAATATAG 245

RESULT 5
US-09-134-000C-1052/C
; Sequence 1052, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; EARLIER FILING DATE: 1998-08-13
; EARLIER APPLICATION NUMBER: US 60/055,778
; EARLIER FILING DATE: 1997-08-15
; EARLIER APPLICATION NUMBER: EPO 93810663.0
; EARLIER FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1052
; LENGTH: 1728
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-1052

Query Match      2.1%; Score 18; DB 4; Length 1728;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 ATAGCAGCAAACTGTATA 669
DB 565 ATAGCAGCAAACTGTATA 548

RESULT 6
US-08-617-785-11
```

; Sequence 11, Application US/08617785E
; Patent No. 6228610
; GENERAL INFORMATION:
; APPLICANT: Flor, Peter J.
; APPLICANT: Kuhn, Ranier
; APPLICANT: Lindaur, Kristen
; APPLICANT: Puttner, Irene
; APPLICANT: Knopfel, Thomas
; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
; FILE REFERENCE: 4-19679/A/PCT
; CURRENT APPLICATION NUMBER: US/08/617,785E
; EARLIER FILING DATE: 1996-03-19
; EARLIER FILING DATE: 1994-09-07
; EARLIER APPLICATION NUMBER: EPO 9416553.7
; EARLIER FILING DATE: 1994-08-19
; EARLIER APPLICATION NUMBER: EPO 93810663.0
; EARLIER FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2745)
US-08-617-785-11

Query Match 2.1%; Score 18; DB 3; Length 2745;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 AACTTCAGCTCAATATAG 349
DB 1502 AACTTCAGCTCAATATAG 1519

RESULT 7
US-09-817-464-11
; Sequence 11, Application US/09817464
; Patent No. 6515107
; GENERAL INFORMATION:
; APPLICANT: Flor, Peter J.
; APPLICANT: Kuhn, Ranier
; APPLICANT: Lindaur, Kristen
; APPLICANT: Puttner, Irene
; APPLICANT: Knopfel, Thomas
; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
; FILE REFERENCE: 4-19679/A/PCT
; CURRENT APPLICATION NUMBER: US/09/817,464
; EARLIER FILING DATE: 2001-03-26
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: EPO 9416553.7
; EARLIER FILING DATE: 1994-08-19
; EARLIER APPLICATION NUMBER: EPO 93810663.0
; EARLIER FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2745)
US-09-817-464-11

Query Match 2.1%; Score 18; DB 4; Length 2745;
Best Local Similarity 100.0%; Pred. No. 23;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 332 AACTTCAGCTCAATATAG 349
DB 1502 AACTTCAGCTCAATATAG 1519

RESULT 8
US-08-617-785-13
; Sequence 13, Application US/08617785E
; Patent No. 6228610
; GENERAL INFORMATION:
; APPLICANT: Flor, Peter J.
; APPLICANT: Kuhn, Ranier
; APPLICANT: Lindaur, Kristen
; APPLICANT: Puttner, Irene
; APPLICANT: Knopfel, Thomas
; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
; FILE REFERENCE: 4-19679/A/PCT
; CURRENT APPLICATION NUMBER: US/08/617,785E
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: EPO 9416553.7
; EARLIER FILING DATE: 1994-09-07
; EARLIER APPLICATION NUMBER: EPO 93810663.0
; EARLIER FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 2766
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2766)
US-08-617-785-13

Query Match 2.1%; Score 18; DB 3; Length 2766;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 AACTTCAGCTCAATATAG 349
DB 1502 AACTTCAGCTCAATATAG 1519

RESULT 9
US-09-817-464-13
; Sequence 13, Application US/09817464
; Patent No. 6515107
; GENERAL INFORMATION:
; APPLICANT: Flor, Peter J.
; APPLICANT: Kuhn, Ranier
; APPLICANT: Lindaur, Kristen
; APPLICANT: Puttner, Irene
; APPLICANT: Knopfel, Thomas
; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
; FILE REFERENCE: 4-19679/A/PCT
; CURRENT APPLICATION NUMBER: US/09/817,464
; EARLIER FILING DATE: 2001-03-26
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: EPO 9416553.7
; EARLIER FILING DATE: 1994-08-19
; EARLIER APPLICATION NUMBER: EPO 93810663.0
; EARLIER FILING DATE: 1993-09-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 2766

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(2766)
US-09-817-464-13

Query Match 2.1%; Score 18; DB 4; Length 2766;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 332 AACCTCAGCTCAATATAG 349
|||||
Db 1502 AACCTCAGCTCAATATAG 1519

RESULT 10

US-08-453-862-1
Sequence 1, Application US/08453862
Patent No. 5738999
GENERAL INFORMATION:

APPLICANT: Segerson, Thomas P.
APPLICANT: Kinzie, J. Mark
APPLICANT: Mulvihill, Eileen R.
APPLICANT: Saugstad, Julie A.
APPLICANT: Westbrook, Gary L.

TITLE OF INVENTION: L-AP4 SENSITIVE GLUTAMATE RECEPTORS

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/453,862

FILING DATE: 30-MAY-1995

CLASSIFICATION: 436

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/176,401

FILING DATE: 30-DEC-1993

ATTORNEY/AGENT INFORMATION:

NAME: Parmelee, Steven W.

REGISTRATION NUMBER: 31,990

REFERENCE/DOCKET NUMBER: 13952-18-2

TELEPHONE: 206-467-9600

TELEFAX: 415-576-0300

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2997 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cdna

FEATURE:

NAME/KEY: CDS

LOCATION: 133..2877

US-08-453-862-1

Query Match 2.1%; Score 18; DB 1; Length 2997;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 332 AACCTCAGCTCAATATAG 349
|||||
Db 1634 AACCTCAGCTCAATATAG 1651

RESULT 11

US-08-452-734A-1
Sequence 1, Application US/08452734A
Patent No. 5831047
GENERAL INFORMATION:

APPLICANT: Segerson, Thomas P.

APPLICANT: Kinzie, J. Mark

APPLICANT: Mulvihill, Eileen R.

APPLICANT: Saugstad, Julie A.

APPLICANT: Westbrook, Gary L.

TITLE OF INVENTION: L-AP4 SENSITIVE GLUTAMATE RECEPTORS

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/452,734A

FILING DATE: 30-MAY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/176,401

FILING DATE: 30-DEC-1993

ATTORNEY/AGENT INFORMATION:

NAME: Parmelee, Steven W.

REGISTRATION NUMBER: 31,990

REFERENCE/DOCKET NUMBER: 13952-18-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-467-9600

TELEFAX: 415-576-0300

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2997 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cdna

FEATURE:

NAME/KEY: CDS

LOCATION: 133..2877

US-08-452-734A-1

Query Match 2.1%; Score 18; DB 2; Length 2997;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 332 AACCTCAGCTCAATATAG 349
|||||
Db 1634 AACCTCAGCTCAATATAG 1651

RESULT 12

US-08-176-401B-1
Sequence 1, Application US/08176401B
Patent No. 6274330
GENERAL INFORMATION:

APPLICANT: Segerson, Thomas P.

APPLICANT: Kinzie, J. Mark

APPLICANT: Mulvihill, Eileen R.

APPLICANT: Saugstad, Julie A.

APPLICANT: Westbrook, Gary L.

TITLE OF INVENTION: L-AP4 SENSITIVE GLUTAMATE RECEPTORS

NUMBER OF SEQUENCES: 5

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Townsend and Crew LLP
;; STREET: Two Embarcadero Center, 8th Floor
;; CITY: San Francisco
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94111-3834
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/176.401B
;; FILING DATE: 30-DECEMBER-1993
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Parmelee, Steven W.
;; REGISTRATION NUMBER: 31,990
;; REFERENCE/DOCKET NUMBER: 13952-18-2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-467-9600
;; TELEFAX: 415-576-0300
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2997 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 133..2877
US-08-176-401B-1

Query Match 2.1%; Score 18; DB 3; Length 2997;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 332 AACTTCAGCTCAATATAG 349
DB 1634 AACTTCAGCTCAATATAG 1651

RESULT 13
PCT-US94-14989-1
; Sequence 1, Application PC/TUS9414989
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: L-AP4 SENSITIVE GLUTAMATE RECEPTORS
; NUMBER OF SEQUENCES: 3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/14989
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176.401
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 13952-18PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2997 base pairs

;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 133..2877
PCT-US94-14989-1

Query Match 2.1%; Score 18; DB 5; Length 2997;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 AACTTCAGCTCAATATAG 349
DB 1634 AACTTCAGCTCAATATAG 1651

RESULT 14
US-09-016-434-1118
; Sequence 1118 Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016.434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1118:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3021 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1370110
US-09-016-434-1118

Query Match 2.1%; Score 18; DB 4; Length 3021;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 AACTTCAGCTCAATATAG 349
DB 1740 AACTTCAGCTCAATATAG 1757

RESULT 15
 US-08-617-785-3
 ; Sequence 3, Application US/08617785E
 ; Patent No. 6228610
 ; GENERAL INFORMATION:
 ; APPLICANT: Flor, Peter J.
 ; APPLICANT: Kuhn, Ranier
 ; APPLICANT: Lindaur, Kristen
 ; APPLICANT: Puttnher, Irene
 ; APPLICANT: Knopfel, Thomas
 ; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
 ; TITLE OF INVENTION: HMR6, HMR7) and Related DNA Compounds
 ; FILE REFERENCE: 4-19679/A/PCT
 ; CURRENT APPLICATION NUMBER: US/08/617,785E
 ; CURRENT FILING DATE: 1996-03-19
 ; EARLIER APPLICATION NUMBER: PCT/EP94/02991
 ; EARLIER FILING DATE: 1994-09-07
 ; EARLIER APPLICATION NUMBER: EPO 9416553.7
 ; EARLIER FILING DATE: 1994-08-19
 ; EARLIER APPLICATION NUMBER: EPO 93810663.0
 ; EARLIER FILING DATE: 1993-09-20
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 3804
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(2604)
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (3325)..(3495)
 ; OTHER INFORMATION: nucleotides designated as n could be a or g or c
 ; OTHER INFORMATION: or t/u
 US-08-617-785-3

Query Match 2.1%; Score 18; DB 3; Length 3804;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 AACCTCAGCTCAATATAG 349
 DB 1337 AACCTCAGCTCAATATAG 1354

Search completed: June 7, 2004, 23:13:46
 Job time: 105 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 21:25:06 ; Search time 417 Seconds
(without alignments)
9386.553 Million cell updates/sec

Title: US-09-889-733b-1

Perfect score: 858

Sequence: 1 atggccgctggccccc.....acaagaagtagactcctga 858

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2995336 seqs, 2280998010 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5991872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	772	90.0	1633	17 US-10-450-826-9	Sequence 9, Appli
2	772	90.0	4208	17 US-10-468-125-22	Sequence 22, Appl
3	610	71.1	3651	9 US-09-925-300-512	Sequence 512, App
4	585	68.2	3778	15 US-10-252-157-328	Sequence 328, App
5	148	17.2	83946	17 US-10-450-826-8	Sequence 8, Appli
6	108	12.6	3189	10 US-09-814-353-20553	Sequence 20553, A
7	60	7.0	60	10 US-09-908-975-9030	Sequence 9030, Ap
8	54	6.3	3008	15 US-10-205-219-83	Sequence 83, Appl
9	33	3.8	332	9 US-09-783-590-5309	Sequence 5309, Ap
10	23	2.7	2523	15 US-10-246-785-16	Sequence 16, Appl
11	23	2.7	2771	10 US-09-361-652-4	Sequence 4, Appli
12	23	2.7	2771	10 US-09-527-315-4	Sequence 4, Appli
13	23	2.7	2771	15 US-10-190-417-4	Sequence 4, Appli
14	20	2.3	21470	9 US-09-764-847-1157	Sequence 1157, Ap

15	20	2.3	21470	15	US-10-092-154-1157	Sequence 1157, Ap
16	20	2.3	1691139	15	US-10-067-514-1	Sequence 1, Appli
17	20	2.3	1691139	16	US-10-419-723-1	Sequence 1, Appli
18	19	2.2	368	9	US-09-764-847-1285	Sequence 1285, Ap
19	19	2.2	368	15	US-10-092-154-1285	Sequence 1285, Ap
20	19	2.2	783	15	US-10-156-761-1744	Sequence 1744, Ap
21	19	2.2	783	13	US-10-282-122A-9827	Sequence 9827, Ap
22	19	2.2	1065	16	US-10-369-493-43191	Sequence 43191, A
23	19	2.2	2033	13	US-10-425-114-11394	Sequence 11394, A
24	19	2.2	2056	13	US-10-424-599-117586	Sequence 117586, A
25	19	2.2	4371	10	US-09-814-353-21999	Sequence 21999, A
26	19	2.2	4412	10	US-09-927-827-43	Sequence 43, Appli
27	19	2.2	7263	15	US-10-037-370-55	Sequence 55, Appli
28	19	2.2	7263	16	US-10-117-722-55	Sequence 7783, Ap
29	19	2.2	8148	10	US-09-764-891-7783	Sequence 1474, Ap
30	19	2.2	47753	13	US-10-087-192-1474	Sequence 174961, A
31	19	2.2	3186778	13	US-10-027-632-174961	Sequence 174961, A
32	19	2.2	3186778	16	US-10-027-632-174961	Sequence 1, Appli
33	19	2.2	9025608	15	US-10-156-761-1	Sequence 1107, Ap
34	18	2.1	269	14	US-10-040-739-1107	Sequence 611, App
35	18	2.1	280	9	US-09-867-550-611	Sequence 247, App
36	18	2.1	435	9	US-09-920-300A-247	Sequence 429, App
37	18	2.1	435	13	US-10-240-425-429	Sequence 247, App
38	18	2.1	435	14	US-10-033-528-247	Sequence 247, App
39	18	2.1	435	15	US-10-099-926-247	Sequence 4857, Ap
40	18	2.1	442	9	US-09-864-761-4857	Sequence 189822, A
41	18	2.1	480	13	US-10-027-632-189822	Sequence 189822, A
42	18	2.1	480	16	US-10-027-632-189822	Sequence 125, App
43	18	2.1	496	9	US-09-960-253-125	Sequence 34714, A
44	18	2.1	527	13	US-10-027-632-34714	Sequence 34714, A
45	18	2.1	527	16	US-10-027-632-34714	

ALIGNMENTS

RESULT 1

US-10-450-826-9
; Sequence 9, Application US/10450826
; Publication No. US20040101818A1
; GENERAL INFORMATION:
; APPLICANT: JI, Darren
; APPLICANT: Axelrod, Douglas W.
; APPLICANT: Cook, Jonathan S.
; APPLICANT: Jaiswal, Neelam
; APPLICANT: Eistein, Richard
; APPLICANT: Houghton, Adam
; APPLICANT: Mertz, Lawrence
; TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiation
; FILE REFERENCE: 044921-5039-WO
; CURRENT APPLICATION NUMBER: US/10/450,826
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 60/255,882
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/285,691
; PRIOR FILING DATE: 2001-04-24
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1633
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. AB001915
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1633)
; OTHER INFORMATION: n = a o r g o r t
US-10-450-826-9

Query Match 90.0%; Score 772; DB 17; Length 1633;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 822; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

201 CGACGAGAGCCTTCCGGACGTGGTCTTCGTGGAGGACGTGGCCGTGGTGTGGAGGAGAC 260

61 CGACGAGAGCCTTCCGGACGTGGTCTTCGTGGAGGACGTGGCCGTGGTGTGGAGGAGAC 120

261 GGCCTTCATCACCCGACCCGGGGCGCGAGCCGGAGGAGGAGTTGACATGATGAAGA 320

121 GGCCTTCATCACCCGACCCGGGGCGCGAGCCGGAGGAGGAGTTGACATGATGAAGA 180

321 AGCATTTAGAAAACTTCAGCTCAATATAGTAGAGATGAAGAATGAAAAATGCAACTTTAGA 380

181 AGCATTTAGAAAACTTCAGCTCAATATAGTAGAGATGAAGAATGAAAAATGCAACTTTAGA 240

381 TGCGCGAGATGTTTTATTACAGCGGAGAGATTTTTTTGGGGCTTTTCCAAAAGGACAAA 440

241 TGCGCGAGATGTTTTATTACAGCGGAGAGATTTTTTTGGGGCTTTTCCAAAAGGACAAA 300

441 T-CAACGAGGTGCTGAAATCTTCGGCTGATACTTTTAAGCACTATGCAAGTCTCCACAGTGC 499

301 TCCACGAGGTGCTGAAATCTTCGGCTGATACTTTTAAGCACTATGCAAGTCTCCACAGTGC 360

500 CAGTGGCAGATGGTTGCATTTCAAGAGTTTCTGCGAGCATGGCTGGGGCCTTAACCTGATCG 559

361 CAGTGGCAGATGGTTGCATTTGAAGAGTTTCTGCGAGCATGGCTGGGGCCTTAACCTGATCG 420

560 CAATTGGGCTTAGTGAATCTGACAGAAAGCCCTTAAAGATCATGCAACAGATGAGTGACC 619

421 CAATTGGGCTTAGTGAATCTGACAGAAAGCCCTTAAAGATCATGCAACAGATGAGTGACC 480

620 ACCGCTACGACAAACTCAGTCTGCTGATGACATAGCAGCAAACTGTATATATCTTAATA 679

481 ACCGCTACGACAAACTCAGTCTGCTGATGACATAGCAGCAAACTGTATATATCTTAATA 540

680 TCCCCAAACAAAGGGCAGCTCTTGTGTCACCGAACCCCGGAGAGATGCCAGAAAGTGCAA 739

541 TCCCCAAACAAAGGGCAGCTCTTGTGTCACCGAACCCCGGAGAGATGCCAGAAAGTGCAA 600

740 AGGTTTATGAGAACTGAAGGACATATGCTGATCCCCGTGAGCATGCTGAACTGGAAA 799

601 AGGTTTATGAGAACTGAAGGACATATGCTGATCCCCGTGAGCATGCTGAACTGGAAA 660

800 AGGTGGATGGCTGCTCAACCTGCTGCTCACTTTTAAATTAACAAGAA 845

661 AGGTGGATGGCTGCTCAACCTGCTGCTCACTTTTAAATTAACAAGAA 706

RESULT 5

US-10-450-826-8/c

Sequence 8, Application US/10450826

Publication No. US2004010181A1

GENERAL INFORMATION:

APPLICANT: Ji, Darren

APPLICANT: Axelrod, Douglas W.

APPLICANT: Cook, Jonathan S.

APPLICANT: Jaiswal, Neelam

APPLICANT: Einstein, Richard

APPLICANT: Houghton, Adam

APPLICANT: Mertz, Lawrence

TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiation

FILE REFERENCE: 044921-5039-WO

CURRENT APPLICATION NUMBER: US/10/450,826

CURRENT FILING DATE: 2003-06-18

PRIOR APPLICATION NUMBER: US 60/255,882

PRIOR FILING DATE: 2000-12-16

PRIOR APPLICATION NUMBER: US 60/285,691

PRIOR FILING DATE: 2001-04-24

NUMBER OF SEQ ID NOS: 149

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 8

LENGTH: 83946

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Genbank Accession No. AL078459

```

US-10-450-826-8

Query Match 17.2%; Score 148; DB 17; Length 83946;
Best Local Similarity 100.0%; Pred. No. 2.6e-66;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 596 AGATCATGCAACAGATGAGTGACCAACCCCTACGACAAACTCACTGTGCTGTGATGACATAG 655
DB 18612 AGATCATGCAACAGATGAGTGACCAACCCCTACGACAAACTCACTGTGCTGTGATGACATAG 18553
QY 656 CAGCAAACTGTATATATCTAAATATCCCAACAAAGGCAGCTTGTGTGACCGAACCC 715
DB 18552 CAGCAAACTGTATATCTAAATATCCCAACAAAGGCAGCTTGTGTGACCGAACCC 18493
QY 716 CGGAGAGGTATCCAGAAAGTGCAAAAGGT 743
DB 18492 CGGAGAGGTATCCAGAAAGTGCAAAAGGT 18465

RESULT 6
US-09-814-353-20553/c
; Sequence 20553, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-005B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20553
; LENGTH: 3189
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3189
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-20553

```

Query Match	12.6%;	Score 108;	DB 10;	Length 3189;
Best Local Similarity	100.0%;	Pred. No. 6.4e-47;		
Matches 108;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	738	AAAGGTTTATGAAACTGAAGACCATATGCTGATCCCGTGACGATGCTCTGAAC	TGGA	797
Db	3187	AAAGGTTTATGAAAACTGAAGACCATATGCTGATCCCGTGACGATGCTCTGAAC	TGGA	3128
Qy	798	AAAGGTGGATGGCGCTCACCTGCTGCTCACTTTTAAATTAACAAGAA		845
Db	3127	AAAGGTGGATGGCGCTCACCTGCTGCTCACTTTTAAATTAACAAGAA		3080
RESULT 7				
US-09-908-975-9030				
; Sequence 9030, Application US/09908975				

Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 9030
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-9030

Query Match 7.0%; Score 60; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 3e-21;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 751 AAATGAGGACCATATGCTGATCCCGTGAGCATGCTGAACTGAAAGGTTGGTGG 810
Db 1 AAATGAGGACCATATGCTGATCCCGTGAGCATGCTGAACTGAAAGGTTGGTGG 60

RESULT 8
US-10-205-219-83
; Sequence 83, Application US/10205219
; Publication No. US2003019803A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbark, Robert
; APPLICANT: Pincock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018200
; CURRENT APPLICATION NUMBER: US/10/205,219
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 83
; LENGTH: 3008
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: N-G, N-G dimethylarginine dimethylaminohydrolase
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 2411, 2412
; OTHER INFORMATION: n is a or g or c or t
US-10-205-219-83

Query Match 6.3%; Score 54; DB 15; Length 3008;
Best Local Similarity 100.0%; Pred. No. 5e-18;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 136 CAGCACCAGCTCTAGTGGGGTGTCTGGGCGAGCAAGCTGGGGCTGCAGGTGGTG 189
Db 567 CAGCACCAGCTCTAGTGGGGTGTCTGGGCGAGCAAGCTGGGGCTGCAGGTGGTG 620

RESULT 9

US-09-783-590-5309
; Sequence 5309, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 5309
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (22)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (42)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (64)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (98)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (134)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (139)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (232)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (257)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (263)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (270)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (272)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (312)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (319)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (326)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-5309

Query Match 3.8%; Score 33; DB 9; Length 332;
Best Local Similarity 100.0%; Pred. No. 8.4e-07;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 441 TCACGAGGTGCTGAATCTCGCTGACTTT 473

Db 65 TCACGAGGCTGCTGAATCTTGGCTGACTTT 97
|||||

RESULT 10
US-10-246-785-16
; Sequence 16, Application US/10246785
; Publication NO. US2003014848A1
; GENERAL INFORMATION:
; APPLICANT: IRM, LLC
; APPLICANT: The Scripps Research Institute
; APPLICANT: Liao, Jiayu
; APPLICANT: Sheng, Ding
; APPLICANT: Schultz, Peter G
; TITLE OF INVENTION: Sweet Taste Receptors
; FILE REFERENCE: 36-002810US/PC
; CURRENT APPLICATION NUMBER: US/10/246,785
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/323,450
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-10-246-785-16

Query Match 2.7%; Score 23; DB 15; Length 2523;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 807 TGGGCTGCTCACCTGCTGCTCAG 829
Db 10 TGGGCTGCTCACCTGCTGCTCAG 32
|||||

RESULT 11
US-09-361-652-4
; Sequence 4, Application US/09361652
; Publication NO. US20030036630A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeier, Juergen
; APPLICANT: Ryba, Nick
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
; FILE REFERENCE: 02307E-088610US
; CURRENT APPLICATION NUMBER: US/09/361,652
; CURRENT FILING DATE: 1999-07-27
; EARLIER APPLICATION NUMBER: US 60/094,465
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2771
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; OTHER INFORMATION: rat G-protein coupled receptor B3 (GPCR-B3)
US-09-361-652-4

Query Match 2.7%; Score 23; DB 10; Length 2771;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 807 TGGGCTGCTCACCTGCTGCTCAG 829
Db 66 TGGGCTGCTCACCTGCTGCTCAG 88
|||||

RESULT 12
US-09-927-315-4
; Sequence 4, Application US/09927315
; Publication NO. US20030040045A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Ryba, Nicholas J.P.
; APPLICANT: Nelson, Greg
; APPLICANT: Hoon, Mark A.
; APPLICANT: Chandrashekar, Jayaram
; APPLICANT: Zhang, Yifeng
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Mammalian Sweet Taste Receptors
; FILE REFERENCE: 02307E-120110US
; CURRENT APPLICATION NUMBER: US/09/927,315
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/302,898
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2771
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; OTHER INFORMATION: rat T1R1 sweet taste receptor
US-09-927-315-4

Query Match 2.7%; Score 23; DB 10; Length 2771;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 807 TGGGCTGCTCACCTGCTGCTCAG 829
Db 66 TGGGCTGCTCACCTGCTGCTCAG 88
|||||

RESULT 13
US-10-190-417-4
; Sequence 4, Application US/10190417
; Publication NO. US20030166137A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Ryba, Nicholas J. P.
; APPLICANT: Chandrashekar, Jayaram
; APPLICANT: Hoon, Mark A.
; APPLICANT: Nelson, Greg
; APPLICANT: Zhang, Yifeng
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: 02307E-120130US
; CURRENT APPLICATION NUMBER: US/10/190,417
; CURRENT FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: US 60/302,898
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 09/927,315
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/358,925
; PRIOR FILING DATE: 2002-02-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2771
; TYPE: DNA
; ORGANISM: Rattus norvegicus
;

; FEATURE:
; OTHER INFORMATION: rat t1r1
US-10-190-417-4

Query Match 2.7%; Score 23; DB 15; Length 2771;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 807 TGGGCTGCTCACCTGCTGCTCAG 829
Db 66 TGGGCTGCTCACCTGCTGCTCAG 88

RESULT 14

US-09-764-847-1157
; Sequence 1157, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1157
; LENGTH: 21470
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1157

Query Match 2.3%; Score 20; DB 9; Length 21470;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 167 GCAAGCTGGGGCTGCAGGTG 186
Db 509 GCAAGCTGGGGCTGCAGGTG 528

RESULT 15

US-10-092-154-1157
; Sequence 1157, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1157
; LENGTH: 21470
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-092-154-1157

Query Match 2.3%; Score 20; DB 15; Length 21470;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 167 GCAAGCTGGGGCTGCAGGTG 186
Db 509 GCAAGCTGGGGCTGCAGGTG 528

Search completed: June 7, 2004, 23:21:00
Job time : 429 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 7, 2004, 21:31:51 ; Search time 378 seconds
(without alignments)
3203.006 Million cell updates/sec

Title: US-09-889-733b-2

Perfect score: 1454

Sequence: 1 MAGLHPSAFGRATHAVRA.....EKVDGLTCCSLINKKVD 285

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool.p/US09889733/runat_07062004_083527_7960/app_query.fasta_1.455
-DB=N_Geneseq_29Jan04 -CFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -IOPC=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09889733@cgn 1.1.708@runat_07062004_083527_7960 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002s:*
7: Geneseq2003as:*
8: Geneseq2003bs:*
9: Geneseq2003cs:*
10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1454	100.0	858	3	AAA47655 Dimethyla
2	1439	99.0	858	9	ADE71465 Human Thr
3	1439	99.0	1633	4	AAH02907 Human she
4	1439	99.0	1633	6	ABQ88102 Human ost
5	1439	99.0	3932	7	ABZ35939 Human sec
6	1439	99.0	4208	6	AAL1484 Drug meta
7	1433	98.6	858	9	ADE71463 Human Met
8	1379	94.8	3008	9	ADBS3014 Primary r

9	1379	94.8	3008	9	ADB79843	AdB79843	Rat N-G,N
10	1200	82.5	3778	9	ADBS3981	Human pro	
11	1064	73.2	3651	3	AAF16077	Human pro	
12	694	47.7	1228	3	AAO06677	Human Ryt	
13	694	47.7	1376	2	AAZ52967	Human pro	
14	688	47.3	858	3	AAA47656	Dimethyla	
15	669.5	46.0	845	3	AAA99390	Human mel	
16	551.5	37.9	770	6	AB576755	Frog embr	
17	482	33.1	873	4	ABL03335	Drosophil	
18	427.5	29.4	444	6	ABN89945	Mouse glo	
19	423	29.1	3223	4	AAK73698	Human imm	
20	423	29.1	28772	6	ABK83555	Human cdn	
21	386	26.5	472	8	ACH21905	Human adu	
22	382	26.3	452	8	ACH27659	Human adu	
23	373.5	25.7	3643	4	ABL03334	Drosophil	
24	352.5	24.2	655	9	ADD34806	Mouse mit	
25	342	23.5	403	8	ACH32042	Human end	
26	330	22.7	862	6	ABQ47886	Oligonucl	
27	330	22.7	862	6	ABQ47887	Oligonucl	
28	330	22.7	898	6	ABQ41321	Oligonucl	
29	330	22.7	898	6	ABQ41320	Oligonucl	
30	314	21.6	862	6	ABQ47888	Oligonucl	
31	314	21.6	862	6	ABQ47889	Oligonucl	
32	314	21.6	898	6	ABQ41319	Oligonucl	
33	314	21.6	898	6	ABQ41318	Oligonucl	
34	301	20.7	765	3	AAA47658	Dimethyla	
35	286	19.7	777	3	AAA47657	Dimethyla	
36	265.5	18.3	83946	6	ABQ88101	Human ost	
37	245.5	16.9	545	9	ADD34805	Mouse mit	
38	237.5	16.3	504	9	ADBS7724	Toxicity-	
39	233	16.0	239	5	AAF68469	Human lun	
40	233	16.0	239	6	ABK38380	CDNA enco	
41	233	16.0	239	7	ACAL0709	Human lun	
42	233	16.0	239	10	ABX99660	Lung canc	
43	233	16.0	239	10	ADE71675	Human lun	
44	204	14.0	6373	5	ABA14642	Human ner	
45	204	14.0	6381	5	ABA14641	Human ner	

ALIGNMENTS

RESULT 1
AAA47655
ID AAA47655 standard; CDNA; 858 BP.
XX
AC AAA47655;
XX
DT 08-NOV-2000 (first entry)
XX Dimethylarginine dimethylaminohydrolase (DDAH1) coding sequence.
DE Dimethylarginine dimethylaminohydrolase; DDAH; DDAH1; DDAH2;
XX arginine deaminase; hyperlipidemia; renal failure; hypertension;
KW restenosis; atherosclerosis; schizophrenia; multiple sclerosis; cancer;
KW ischemia reperfusion injury; septic shock; multi organ failure;
KW arthritis; skin disorders; inflammatory cardiac disease; migraine;
KW infection; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..858
FT /*tag= a
FT /product= "DDAH1"

XX PA (UNLO) UNIV COLLEGE LONDON.
 XX PI Vallance PUT, Leiper JM, Whitley GSJ, Charles IG;
 XX DR WPI: 2000-543392/49.
 XX PT P-PSDB; AAB01477.
 XX PT Novel methylarginase polypeptides and polynucleotides, used to identify
 PT modulators of them, which are used in the treatment of e.g. cancer,
 PT hypertension, and bacterial infections.
 XX PS Claim 1; Page 55-56; 68pp; English.
 XX CC Nucleotides encoding methylarginase polypeptides, vectors comprising
 CC these nucleotides and the polypeptides themselves can be used in
 CC medicaments for the treatment of hyperlipidemia, renal failure,
 CC hypertension, restenosis after angioplasty, atherosclerosis,
 CC complications of heart failure, schizophrenia, multiple sclerosis or
 CC cancer. Modulators of the enzyme can be used in medicaments for the
 CC treatment of ischemia-reperfusion injury of the brain or heart, cancer,
 CC lethal hypertension in severe inflammatory conditions such as septic
 CC shock or multi-organ failure, or local and systemic inflammatory
 CC disorders including arthritis, skin disorders, inflammatory cardiac
 CC disease, migraine, or microbial or bacterial infection
 XX SQ Sequence 858 BP; 217 A; 214 C; 256 G; 171 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2,52e-155 Length: 858
 Score: 1454.00 Matches: 285
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-09-889-733B-2 (1-285) x AAA47655 (1-858)

QY 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
 Db 1 ATGGCGGGCTCGGCACCCCTTCGCCCTTCGCCCGGGCCACCCAGCCGCTGGTGGCGG 60
 QY 21 LeuProGluSerLeuGlyHisAlaLeuArgSerAlaLysGlyGluValAlaVal 40
 Db 61 CTACCCGAGTCTGCTGCAGCACCGCTGAGAGCGCCAGCGGCGAGGAGTGGACGTC 120
 QY 41 AlaArgAlaGluArgGlnHisGlnLeuValGlyValLeuGlySerLeuGlyLeu 60
 Db 121 GCCCGCGGGAACGGCAGCAGCAGCTCTACGTTGGCGCTGCTGGGCGAGCAAGCTGGGCGTG 180
 QY 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
 Db 181 CAGGTGGTGGAGCTGCGGCGCCGACGAGAGCTTCGGACTCGCTTCGTGGAGGACGTG 240
 QY 81 AlaValValCysGluLeuThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLys 100
 Db 241 GCCGTGGTGGAGGAGACGCGCTTCATCACCCGCGCGCGCGCGCGCGCGCGAGGAG 300
 QY 101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnLeuValGluMetLys 120
 Db 301 GAGGTTCATGATGAGAGACGATAGAAACCTTCAGCTCAATATATAGTAGAGATGAAA 360
 QY 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheVal 140
 Db 361 GATGAAATGCAACTTAGATGGCGAGATGTTTATTACAGGCGAGAGAAATTTTTTGTG 420
 QY 141 GlyLeuSerLeuArgThrAsnGlnArgGlyAlaGluLeuLeuAlaAspThrPheLeuAsp 160
 Db 421 GGCTTTTCCAAAGACAAATCAAGAGGTGCTGAAATCTTGCTGATATTTTAAAGAC 480
 QY 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
 Db 481 TATGAGTCTCCACAGTGCCAGTGGCAGATGGGTTGCAATTTGAAGAGTTTCTGACGATG 540

QY 181 AlaGlyProAsnLeuLeuAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysile 200
 Db 541 GCTGGGCTTAACCTGATCGCAATGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATC 600
 QY 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
 Db 601 ATGCAACAGATGAGTACCAACCGCTACGACAAACTCACTGTGCTGATGACATAGCAGCA 660
 QY 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
 Db 661 AACTGTATATATCTAATATATCCCAACAAGGCGAGCTTCTGTCACCGAACCCCGGAA 720
 QY 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuLeuProVal 260
 Db 721 GAGTATCCCAAGAGTCAAAAGTTTATGAGAACTGAGAGACCAATATGCTGATCCCGTG 780
 QY 261 SerMetSerGluLeuGluLysValLeuGlyLeuThrCysCysSerValLeuLeuAsn 280
 Db 781 AGCATGCTGAACTGGAAAGGTGGATGGCTGCTCACCTGCTGCTCAGTTTATTAATAC 840
 QY 281 LysLysValAspSer 285
 Db 841 AAGAAGGTAGACTCC 855
 RESULT 2
 ADE71465
 ID ADE71465 standard; DNA; 858 BP.
 XX AC ADE71465;
 XX AC ADE71465;
 DT 29-JAN-2004 (first entry)
 XX DE Human Thr87 wild-type dimethylarginine dimethylaminohydrolase-1 DNA.
 XX KW Human; dimethylarginine dimethylaminohydrolase-1; DDAH 1; chromosome 1;
 KW single nucleotide polymorphism; SNP; ADMA;
 KW asymmetrical dimethyl-arginine; nitric oxide; cardiovascular disease;
 KW coronary heart disease; cerebrovascular disease; hypertension; diabetes;
 KW susceptibility; genotyping; therapeutic targeting; antidiabetic;
 KW cardiant; cerebroprotective; hypotensive; gene therapy; gene; ds.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX CDS 1..858
 FT /tag= a
 FT /product= "wild-type (Thr87) DDAH 1"
 FT replace(260,T)
 FT /tag= b
 FT /standard name= "Single nucleotide polymorphism (SNP)"
 FT /note= "Results in Thr (C variant) or Met (T variant) at
 position 87 of the protein"
 XX PN W02003089638-A1.
 XX 30-OCT-2003.
 XX 11-APR-2003; 2003WO-FI000274.
 XX 19-APR-2002; 2002US-00125456.
 XX (JURI-) JURILAB LTD OY.
 XX Valkonen V, Salonen JT, Pirkkanen M, Tuomainen T, Laakso J;
 XX Laaksonen R;
 XX WPI: 2003-854121/79.
 XX P-PSDB; ADE71465.
 XX New nucleic acid encoding a variant dimethylarginine
 PT dimethylaminohydrolase-1 (DDAH-1) protein, useful for treating diabetes,
 PT and its vascular complications, e.g. coronary or cerebrovascular disease

XX Claim 20; Page 404-407; 678pp; Japanese.

XX The present invention provides the protein and coding sequences of a

XX number of human shear stress response proteins. These are useful in the

CC diagnosis, treatment and screening of vascular diseases caused by

CC arteriosclerosis, including heart failure, post-PTCA restenosis and

CC hypertension

XX SQ Sequence 1633 BP; 400 A; 406 C; 432 G; 393 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 3,22e-153 Length: 1633

Score: 1439.00 Matches: 283

Percent Similarity: 99.65% Conservativity: 1

Best Local Similarity: 99.30% Mismatches: 0

Query Match: 98.97% Indels: 0

DB: 4 Gaps: 0

US-09-889-733B-2 (1-285) x AAH02907 (1-1633)

QY 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20

DB 323 ATGGCGGGCTCGGCGCCCGGCGCTTCGGCGGGCCACCCACCGCGGTGGTGGCGG 382

QY 21 LeuProGluSerLeuGlyGlnHisAlaLeuArgSerAlaLysGlyGluGluValAspVal 40

DB 383 CTACCCGAGTGGCTCGGCGCAGCAGCGCTGAGAAAGCGCCAAAGGGCGAGAGGTGGACGTC 442

QY 41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60

DB 443 GCCCGCGGGAACGCGACACAGCTCTACGTGGCGTGTGGCAGCAGCTGGGCTG 502

QY 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80

DB 503 CAGGTGGTGGAGTCCGCGCGCAGCAGAGCCTTCGGACTCGCTCTTCGTGGAGGAGCGTG 562

QY 81 AlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLys 100

DB 563 GCGGTGGTGGAGGAGCGCGCTCTACCCAGCCCGGGCGCGCGAGCGAGGAG 622

QY 101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120

DB 623 GAGGTGGATCATGATGAAAGAGCATTAAGAAACCTTCAGCTCAATATAGTAGATGAAA 682

QY 121 AspGluAsnAlaThrLeuAspGlyCysValValLeuPheThrGlyArgGluPheVal 140

DB 683 GATGAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGAGAGATTTTGTG 742

QY 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp 160

DB 743 GACCTTTCCAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACATTTAAGAC 802

QY 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180

DB 803 TATGAGTCTCCACAGTCCAGTGGCAGATGGTGTGCAATTTGAAGAGTTTCTGAGCATG 862

QY 181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200

DB 863 GTGGGCTTAACCTGATCGCAATTCGGTCTAGTGAATCTGCACAGAGCCCTTAAGATC 922

QY 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220

DB 923 ATGCAACAGATGAGTGACACCGCTACGCAAACTCACTGTGCCTGATGATGATGACAGCA 982

QY 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240

DB 983 AACTGTATATATCTAATATATCCCAACAAAGGCGACGCTCTTGTGCGACCGACCCCGGAA 1042

QY 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal 260

DB 1043 GAGTATCCAGAAAGTGCAAGGTTTATGAGAAACTTAAGAGGACCATATCTGATCCCGGTG 1102

QY 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuLeuLeu 280

DB 1103 AGCATGCTCTGAACCTGAAAAGGTGGATGGGTGCTCACTGCTGCTCAGTTTAAATTAAC 1162

QY 281 LysLysValAspSer 285

DB 1163 AAGAAAGTAGACTCC 1177

RESULT 4

ABQ88102

ID ABQ88102 standard; cDNA; 1633 BP.

XX AC ABQ88102;

XX DT 18-SEP-2002 (first entry)

XX DE Human osteoblast differentiation related cDNA SEQ ID NO 9.

XX KW Human; osteoblast; stem cell differentiation; bone tissue deposition;

XX KW osteoporosis; osteopathic; ss.

XX OS Homo sapiens.

XX PN WO200250301-A2.

XX PD 27-JUN-2002.

XX PF 18-DEC-2001; 2001WO-US048276.

XX PR 18-DEC-2000; 2000US-0255882P.

XX PR 24-APR-2001; 2001US-0285691P.

XX PA (GENE-) GENE LOGIC INC.

XX PA (PROC) PROCTER & GAMBLE CO.

XX PI Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;

XX PI Mertz L;

XX PI WPI; 2002-557663/59.

XX PT Use of genes and their expression profiles associated with osteoblast

XX PT differentiation for screening modulators bone formation, for diagnosing

XX PT or treating e.g. osteoporosis, or as markers for the differentiation

XX PS process.

XX PS Claim 1; SEQ ID NO 9; 78pp + Sequence Listing; English.

XX PS The invention relates to genes and their expression profiles are used

CC for: (a) screening modulators of precursor stem cell differentiation into

CC osteoblasts, or bone tissue deposition; (b) diagnosing abnormal

CC deposition of bone tissue, abnormal rate of osteoblast formation or

CC osteoporosis; or (c) treating or monitoring treatment of the conditions

CC cited in (b), or monitoring the progression of bone tissue deposition.

CC Specific conditions include postmenopausal osteoporosis, glucocorticoid

CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-

CC induced abnormalities in bone formation or bone loss, conditions that

CC involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),

CC skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome

CC or fibrous dysplasia. The present sequence is that of an osteoblast

CC differentiation associated cDNA marker of the invention. Note: The

CC sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1633 BP; 400 A; 406 C; 432 G; 393 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 3,22e-153 Length: 1633

Score: 1439.00 Matches: 283

Percent Similarity: 99.65% Conservativity: 1

Best Local Similarity: 99.30% Mismatches: 0

Query Match: 98.97% Indels: 0

DB: 4 Gaps: 0

XX Claim 20; Page 404-407; 678pp; Japanese.

XX The present invention provides the protein and coding sequences of a

XX number of human shear stress response proteins. These are useful in the

CC diagnosis, treatment and screening of vascular diseases caused by

CC arteriosclerosis, including heart failure, post-PTCA restenosis and

CC hypertension

XX SQ Sequence 1633 BP; 400 A; 406 C; 432 G; 393 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 3,22e-153 Length: 1633

Score: 1439.00 Matches: 283

Percent Similarity: 99.65% Conservativity: 1

Best Local Similarity: 99.30% Mismatches: 0

Query Match: 98.97% Indels: 0

DB: 4 Gaps: 0

US-09-889-733B-2 (1-285) x AAH02907 (1-1633)

QY 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20

DB 323 ATGGCGGGCTCGGCGCCCGGCGCTTCGGCGGGCCACCCACCGCGGTGGTGGCGG 382

QY 21 LeuProGluSerLeuGlyGlnHisAlaLeuArgSerAlaLysGlyGluGluValAspVal 40

DB 383 CTACCCGAGTGGCTCGGCGCAGCAGCGCTGAGAAAGCGCCAAAGGGCGAGAGGTGGACGTC 442

QY 41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60

DB 443 GCCCGCGGGAACGCGACACAGCTCTACGTGGCGTGTGGCAGCAGCTGGGCTG 502

QY 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80

DB 503 CAGGTGGTGGAGTCCGCGCGCAGCAGAGCCTTCGGACTCGCTCTTCGTGGAGGAGCGTG 562

QY 81 AlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLys 100

DB 563 GCGGTGGTGGAGGAGCGCGCTCTACCCAGCCCGGGCGCGCGAGCGAGGAG 622

QY 101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120

DB 623 GAGGTGGATCATGATGAAAGAGCATTAAGAAACCTTCAGCTCAATATAGTAGATGAAA 682

QY 121 AspGluAsnAlaThrLeuAspGlyCysValValLeuPheThrGlyArgGluPheVal 140

DB 683 GATGAAATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGAGAGATTTTGTG 742

QY 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp 160

DB 743 GACCTTTCCAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATACATTTAAGAC 802

QY 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180

DB 803 TATGAGTCTCCACAGTCCAGTGGCAGATGGTGTGCAATTTGAAGAGTTTCTGAGCATG 862

QY 181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200

DB 863 GTGGGCTTAACCTGATCGCAATTCGGTCTAGTGAATCTGCACAGAGCCCTTAAGATC 922

QY 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220

DB 923 ATGCAACAGATGAGTGACACCGCTACGCAAACTCACTGTGCCTGATGATGATGACAGCA 982

QY 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240

DB 983 AACTGTATATATCTAATATATCCCAACAAAGGCGACGCTCTTGTGCGACCGACCCCGGAA 1042

QY 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal 260

DB 1043 GAGTATCCAGAAAGTGCAAGGTTTATGAGAAACTTAAGAGGACCATATCTGATCCCGGTG 1102

QY 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuLeuLeu 280

DB 1103 AGCATGCTCTGAACCTGAAAAGGTGGATGGGTGCTCACTGCTGCTCAGTTTAAATTAAC 1162

QY 281 LysLysValAspSer 285

DB 1163 AAGAAAGTAGACTCC 1177

RESULT 4

ABQ88102

ID ABQ88102 standard; cDNA; 1633 BP.

XX AC ABQ88102;

XX DT 18-SEP-2002 (first entry)

XX DE Human osteoblast differentiation related cDNA SEQ ID NO 9.

XX KW Human; osteoblast; stem cell differentiation; bone tissue deposition;

XX KW osteoporosis; osteopathic; ss.

XX OS Homo sapiens.

XX PN WO200250301-A2.

XX PD 27-JUN-2002.

XX PF 18-DEC-2001; 2001WO-US048276.

XX PR 18-DEC-2000; 2000US-0255882P.

XX PR 24-APR-2001; 2001US-0285691P.

XX PA (GENE-) GENE LOGIC INC.

XX PA (PROC) PROCTER & GAMBLE CO.

XX PI Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;

XX PI Mertz L;

XX PI WPI; 2002-557663/59.

XX PT Use of genes and their expression profiles associated with osteoblast

XX PT differentiation for screening modulators bone formation, for diagnosing

XX PT or treating e.g. osteoporosis, or as markers for the differentiation

XX PS process.

XX PS Claim 1; SEQ ID NO 9; 78pp + Sequence Listing; English.

XX PS The invention relates to genes and their expression profiles are used

CC for: (a) screening modulators of precursor stem cell differentiation into

CC osteoblasts, or bone tissue deposition; (b) diagnosing abnormal

CC deposition of bone tissue, abnormal rate of osteoblast formation or

CC osteoporosis; or (c) treating or monitoring treatment of the conditions

CC cited in (b), or monitoring the progression of bone tissue deposition.

CC Specific conditions include postmenopausal osteoporosis, glucocorticoid

CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-

CC induced abnormalities in bone formation or bone loss, conditions that

CC involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),

CC skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome

CC or fibrous dysplasia. The present sequence is that of an osteoblast

CC differentiation associated cDNA marker of the invention. Note: The

CC sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1633 BP; 400 A; 406 C; 432 G; 393 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 3,22e-153 Length: 1633

Score: 1439.00 Matches: 283

Percent Similarity: 99.65% Conservativity: 1

Best Local Similarity: 99.30% Mismatches: 0

Query Match: 98.97% Indels: 0

DB: 4 Gaps: 0

US-09-889-733B-2 (1-285) x AB088102 (1-1633)

QY 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20

Db 323 ATGCCGGGCTCGGCCACCCCGCGCTTCGGCGGGCCACCCACCCCTGTGTGGGGC 382

QY 21 LeuProGluSerLeuGlyHisAlaLeuArgSerAlaGlyGlyGluValVal 40

Db 383 CTACCCGAGTCGTCGCCACGACGCGCTGAGAAAGCCCAAGGCGGAGGAGTGGACGTC 442

QY 41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLeuGlyLeu 60

Db 443 GCCCGCGGAGCGGACGACGACCCAGCTCTACGTGGCGCTGTGGCGAGCAAGCTGGGGCTG 502

QY 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80

Db 503 CAGGTGGTGGAGTCGCGCGCGACGAGAGCTTCGCGACTGCGTCTTCGTGGAGACGTCG 562

QY 81 AlaValValCysGluGlnThrAlaLeuLeuThrArgProGlyAlaProSerArgArgGly 100

Db 563 GCCGTGGTGGCGAGAGCGCGCTCATCCCGGACCCCGCGCGCGCGAGCGAGGAG 622

QY 101 GluValAspMetMetLeuGlnAlaLeuGluLysLeuGlnLeuAsnLeuValGluMetLys 120

Db 623 GAGGTTGACATGATGAAGAGGACGATTAGAAAACCTTCAGCTCAATATAGTAGAGTAA 582

QY 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal 140

Db 683 GATGAAATGCACTTTAGATGCGGAGATGTTTATTACACGCGGAGAGAAATTTTGTG 742

QY 141 GlyLeuSerLeuArgThrGlnArgGlyAlaGluLeuAlaAspThrPheLysAsp 160

Db 743 GGCCTTTCAAAGGACAAATCAACGAGGTCTGAAATCTTGCGTATCTTTTAAAGAC 802

QY 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180

Db 803 TATGAGCTTCCACAGTCGCGAGTGGCGAGATGGTTGCAATTTGAAGAGTTTTCGACGATG 862

QY 181 AlaGlyProAsnLeuLeuAlaGlySerSerGluSerAlaGlnLysAlaLeuLysLeu 200

Db 863 GCTGGGCGCTAACTGATCGCAATGGTCTAGTGAATCTGCACAGAGCCCTTAAAGATC 922

QY 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspLeuAla 220

Db 923 ATGCAACAGATGATGACCCCGCTACGACAACTCACTGCTGCTGATGACATAGCAGCA 982

QY 221 AsnCysLeuTyrLeuAsnLeuProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240

Db 983 AACTGTATATATCTAAATATCCCAACAAAGGCGACGCTTCTGTCACCGAACCCCGGAA 1042

QY 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuLeuProVal 260

Db 1043 GAGTATCCAGAAAGTCAAGGTTTATGAAACTGAAGACCATATGCTGATCCCGCTG 1102

QY 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuLeuLeu 280

Db 1103 AGCATGCTCTGAATGAAAGGATGGATGGCTGCTCACCTGCTGCTCAGTTTAAATTAAC 1162

QY 281 LysLeuValAspSer 285

Db 1163 AGAAAGTAGACTCC 1177

RESULT 5

ID ABZ35939

XX ABZ35939 standard; cDNA; 3932 BP.

AC ABZ35939;

XX

DT 10-FEB-2003 (first entry)

XX

DE Human secretory polynucleotide SPTM SEQ ID NO 103.

XX

Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia; asthma; Crohn's disease; neurological disorder; epilepsy; cancer; Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease; multiple sclerosis; Parkinson's disease; cell proliferative disorder; anti-inflammatory; immunosuppressive; neuroprotective; nootropic; neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic; anipsoriatic; antianemic; anti-HIV; human immunodeficiency virus; secretory polynucleotide; secretory protein; gene; ss.

Homo sapiens.

WO200283876-A2.

24-OCT-2002.

27-MAR-2002; 2002WO-US009921.

29-MAR-2001; 2001US-0280067P.

29-MAR-2001; 2001US-0280058P.

16-MAY-2001; 2001US-0291280P.

17-MAY-2001; 2001US-0291829P.

17-MAY-2001; 2001US-0291849P.

19-JUN-2001; 2001US-0299428P.

20-JUN-2001; 2001US-0299776P.

20-JUN-2001; 2001US-0300001P.

(INCY-) INCYTE GENOMICS INC.

Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J; Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshay SR; Daugherty SC, Dam TC, Liu TP, Nguyen AJ, Kleinfeld Y, Gerstein EH; Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B; Flores V, Marwaha R, Lo A, Lan RY, Urashka ME; WPI; 2003-075543/07.

P-PSDB; ABP75492.

New human secretory proteins and polynucleotides, useful for diagnosing, treating or preventing autoimmune/inflammatory disorders (e.g. AIDS), neurological disorders (e.g. Alzheimer's), or cell proliferations or cancers.

Claim 1; SEQ ID NO 103; 458bp + Sequence Listing; English.

The invention relates to a secretory polynucleotide (designated sptm) comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a naturally occurring polynucleotide sequence at least 90 % identical to the polynucleotide sequence, a polynucleotide complementary to them or an RNA equivalent of them. The polypeptide or polynucleotide are useful for treating, preventing or diagnosing a disease or condition associated with the expression of functional SPTM. These are particularly useful for diagnosing, treating or preventing autoimmune/inflammatory disorders (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's disease), neurological disorders (e.g. epilepsy, Huntington's disease, dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease, multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety, schizophrenia or amnesia), or cell proliferative disorders (e.g. psoriasis, polycythemia vera, or cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain, breast, cervix or prostate). Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 3932 BP; 1089 A; 883 C; 775 G; 1185 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,13e-152	Length:	3932
Score:	1439.00	Matches:	283
Percent Similarity:	99.65%	Conservative:	1
Best Local Similarity:	99.30%	Mismatches:	1
Query Match:	98.97%	Indels:	0
DB:	7	Gaps:	0

US-09-889-733B-2 (1-285) x ABZ35939 (1-3932)

Qy 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
Db 17 ATGCCCGGGCTCGGCCACCCCGCCCTTCGGCGGGCCACCCACCGCGGGTGGCGGGCG 76
Qy 21 LeuProGluSerLeuGlyHisAlaLeuArgSerAlaGlyGlyGluGluValAspVal 40
Db 77 CTACCCGAGTGGCTCGGCCAGCAGCGCTGGAAGCGCCCAAGGGCGAGAGGTGGAGCTC 136
Qy 41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
Db 137 GCCCGCGCGGAACGCGACAGCAGCTCTAGCTGGCGCTGCTGGCGCAGCAAGCTGGGGCTG 196
Qy 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
Db 197 CAGGTGGTGGAGCTGCGCGCCGAGAGAGCTTCGAGCTTCGGTTCGGTGGAGAGCTG 256
Qy 81 AlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLys 100
Db 257 GCCGTGGTGGAGGAGAGCGCCCTCATCACCGACCCCGGGCGCGAGCGGAGGAAG 316
Qy 101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
Db 317 GAGGTGACATGATGAAGAAGAGCAATTAGAAACCTTCAGCTCAATATAGTAGAGATGAAA 376
Qy 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheVal 140
Db 377 GATGAATGCAACTTATAGTGGCGAGAGTCTTATTCACAGCCACAGATTTTGTG 436
Qy 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp 160
Db 437 GGCCTTTCCAAAGAGCAAAATCAACGAGGTGCTGAAATCTTGGCTGATCTTTTAAGGAC 496
Qy 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLysLysSerPheCysSerMet 180
Db 497 TATGAGTCTCCAGTCCAGTCCAGTGGAGATGGTGTGATTTGAGAGATTTCTGACAGATG 556
Qy 181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200
Db 557 GCTGGCCCTAACCTGATCGCAATGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATC 616
Qy 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
Db 617 ATGCAACAGATGATGACCCCGCTACGACAAATCTACTGTGCTGTATGATCAGACGCA 676
Qy 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
Db 677 AACTGTATATATCTAAATATATCCCAACAAAGGGCACGCTTGTGTCACCGAACCCCGAA 736
Qy 241 GluTyrProGluSerAlaIleValTyrGlnLysLysLysAspHisMetLeuIleProVal 260
Db 737 GAGTATCCAGAAAGTCAAGAGTTTATGAGAACTGAAGGACCATATGCTGATCCCCGTG 796
Qy 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuIleAsn 280
Db 797 AGCATGTCTGAAGTGAAGAGGTGATGGCTGCTCAGCTGCTCAGTTTAAATTAAC 856
Qy 281 LysLysValAspSer 285
Db 857 AAGAAGTAGACTCC 871

RESULT 6

AAL41494

ID AAL41494 standard; DNA; 4208 BP.

XX

AC AAL41494;

XX

DT 21-NOV-2002 (first entry)

XX

DE Drug metabolising enzyme encoding DNA - 7671089CB1.

XX

KW Anti-HIV; antiarteriosclerotic; dermatological; cytostatic; thyromimetic;

osteopathic; thrombolytic; ophthalmological; antilipemic; hepatotropic;
antidiarrheic; antiinflammatory; virucide; immunogenic; autoimmune;
drug metabolising enzyme; DME; inflammatory; AIDS; atherosclerosis;
contact dermatitis; cell-proliferative; cancer; cirrhosis; dwarfism;
developmental; hypothyroidism; endocrine; osteoporosis; thrombosis;
diabetes; glaucoma; keratitis; metabolic; hyperlipidaemia; diarrhoea;
cystic fibrosis; gastrointestinal; gastroenteritis; liver; hepatitis;
Reye's syndrome; exogenous compound; gene therapy; enzyme; human; gene;
ds.

Homo sapiens.

Key Location/Qualifiers

CDS 200..1225

/*tag= a

/product= "Drug metabolising enzyme protein"

WO200266654-A2.

XX 29-AUG-2002.

XX 14-FEB-2002; 2002WO-US004918.

XX 16-FEB-2001; 2001US-0269643P.

XX 23-FEB-2001; 2001US-0271332P.

XX 16-MAR-2001; 2001US-0276767P.

XX 06-APR-2001; 2001US-0282077P.

XX 19-APR-2001; 2001US-0285447P.

XX 27-APR-2001; 2001US-0287060P.

XX 03-MAY-2001; 2001US-0288543P.

XX (INCY-) INCYTE GENOMICS INC.

XX Astromoff A, Au-Young J, Baughn MR, Ding L, Duggan BM;

XX Forsythe IJ, Gietzen KU, Griffin JA, Lee EA, Lu Y, Richardson TW;

XX Ring HZ, Sanjanwala MM, Swarnakar A, Wallia NK, Warren BA, Xu Y;

XX Yue H, Zebajadian Y;

XX WPI; 2002-674949/72.

XX P-PSDB; ARO22798.

XX New drug metabolizing enzymes (DME) useful for diagnosing, treating and

XX preventing diseases or conditions associated with aberrant DME

XX expression, e.g. cancer, AIDS, atherosclerosis, diabetes, glaucoma,

XX hepatitis, osteoporosis.

XX Claim 5; Page 164-165; 166pp; English.

XX The invention relates to an isolated polypeptide comprising 12 fully

XX defined sequences of 81-615 amino acids given in the specification; a

XX naturally occurring amino acid sequence at least 90% identical to, having

XX 81-559 amino acids, at least 93% identical to a sequence of 529 amino

XX acids, or at least 97% identical to a sequence of 615 amino acids, all

XX given in the specification; or a biologically active or immunogenic

XX fragment of the polypeptide. The polypeptides and polynucleotides are

XX useful in diagnosing, treating and preventing diseases or conditions

XX associated with the decreased expression or overexpression of a drug

XX metabolising enzyme (DME), such as autoimmune/inflammatory (e.g. AIDS,

XX atherosclerosis, contact dermatitis) cell-proliferative (e.g. cancer,

XX cirrhosis), developmental (e.g. dwarfism, hypothyroidism), endocrine

XX (e.g. osteoporosis, thrombosis, diabetes), eye (e.g. glaucoma,

XX keratitis), metabolic (e.g. hyperlipidaemia, cystic fibrosis),

XX gastrointestinal (e.g. gastroenteritis, diarrhoea), or liver (e.g.

XX hepatitis, Reye's syndrome) disorders. These are also useful in assessing

XX the effects of exogenous compounds on the expression of nucleic acid and

XX amino acid sequences of DME. The DME or its fragments are useful in

XX screening compounds for effectiveness as agonist or antagonist of the

XX polypeptides, or in altering the expression of the target polynucleotide

XX and compounds that specifically bind to or modulate the activity of the

XX polypeptide. The microarray is useful in monitoring or measuring protein-

XX protein interactions, drug-target interactions, and gene expression

XX profiles. The polynucleotides of the invention can be used in gene

XX therapy. This polynucleotide sequence represents the DNA encoding a drug

CC metabolising enzyme of the invention
SQ Sequence 4208 BP; 1092 A; 1017 C; 883 G; 1216 T; 0 U; 0 Other;

Alignment Scores:
P-Seq. No.: 1.25e-152 Length: 4208
Score: 1439.00 Matches: 283
Percent Similarity: 99.65% Conservative: 1
Best Local Similarity: 99.30% Mismatches: 1
Query Match: 98.9% Indels: 0
DB: 6 Gaps: 0

US-09-889-733B-2 (1-285) x AAL41494 (1-4208)

QY 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
DB 368 ATGGCGGGCTGGCCACCCCGCGCTTCGGCGGCCACCCACCGCGTGGTGGCGGCG 427

QY 21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluGluValAspVal 40
DB 428 CTACCCGAGTCTGGCGCAGCAGCGCTGAGAGCGCAAGGGCGAGAGGTGACGTC 487

QY 41 AlaArgAlaGluArgGlnHisGlnLeuThrValGlyValLeuGlySerLysLeuGlyLeu 60
DB 488 GCCCGCGCGAAGCGGACGACCGCTCTACGTGGCGGTGCTGGCGGACGAGCTGGGGCTG 547

QY 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
DB 548 CAGGTGGTGGAGCTGGCGCGCAGCAGAGCCCTTCGGACTGGCTCTTCGTGGAGGACGTG 607

QY 81 AlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLys 100
DB 608 GCCGTGGTGTGGAGGAGAGCGCCCTTCATCACCACCGCGCGCGCGGAGCGGAGGAG 667

QY 101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
DB 668 GAGGTTCATGATGAAAGAGCAATAGAAACCTTCAGCTCAATATAGTAGAGATGAA 727

QY 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal 140
DB 728 GATGAAATGCAACTTTAGATGCGGAGATGTTTATTACAGCGGACAGAAATTTTGTG 787

QY 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuLeuAlaAspThrPheLysAsp 160
DB 788 GGCCTTTCCAAAGAGCAATCAACAGAGTGTGTAATCTTGGCTGATACTTTTAAGAC 847

QY 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
DB 848 TATGCAAGTCTCCACAGTGGCAGATGGTGGCTTGCATTTGAAGAGTTTCTTGCACATG 907

QY 181 AlaGlyProAsnLeuLeuAlaLeuGlySerGluSerAlaGlnLysAlaLeuLysIle 200
DB 908 GCTGGGCCCTTAACCTGATCGCAATTTGGTCTAGTGAATCTGCACAGAGCCCTTAAGATC 967

QY 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
DB 968 ATGCAACAGATGATGACACCGCTACGACAACTCACCTGCTGCTGATGACATAGCAGCA 1027

QY 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
DB 1028 AACTGTATATATATAATATCCCAACAAAGGGGACGCTTCTGTCACGACCCCGGAA 1087

QY 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal 260
DB 1088 GAGTATCCAGAAAGTCCAAAGGTTTATGAGAACTGAAGGACCATATGCTGATCCCCGTG 1147

QY 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuLeuAsn 280
DB 1148 AGCATGCTTGAATGGAAGGAGGATGGTGGCTGCTCACCTGTGCTGCTGCTTAAATTAAC 1207

QY 281 LysLysValAspSer 285
DB 1208 AGAAAGTAGACTCC 1222

RESULT 7
ADE71463
ID ADE71463 standard; DNA; 858 BP.
XX ADE71463;
XX ADE71463;
XX 29-JAN-2004 (first entry)
DE Human Met87 variant dimethylarginine dimethylaminohydrolase-1 DNA.
XX Human; variant dimethylarginine dimethylaminohydrolase-1; variant DDAH 1;
XX chromosome 1; single nucleotide polymorphism; SNP; ADMA;
XX asymmetrical dimethyl-arginine; nitric oxide; cardiovascular disease;
XX coronary heart disease; cerebrovascular disease; hypertension; diabetes;
XX susceptibility; genotyping; therapeutic targeting; antidiabetic;
XX cardiac; cerebroprotective; hypotensive; gene therapy; gene; ds.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 1..858
FT /*tag= a
FT /product= "Met87 variant DDAH 1"
FT replace(260,C)
FT /*tag= b
FT /standard name= "Single nucleotide polymorphism (SNP)"
FT /note= "Results in Thr (C variant) or Met (T variant) at
FT position 87 of the protein"
XX WO2003089638-A1.
XX 30-OCT-2003.
XX 11-APR-2003; 2003WO-FI000274.
XX 19-APR-2002; 2002US-00125456.
XX (JURI-) JURILAB LTD OY.
XX Valkonen V, Salonen JT, Pirskanen M, Tuomainen T, Laakso J;
XX Laaksonen R;
XX WPI; 2003-854121/79.
XX P-PSDB; ADE71464.
XX New nucleic acid encoding a variant dimethylarginine
XX dimethylaminohydrolase-1 (DDAH-1) protein, useful for treating diabetes,
XX and its vascular complications, e.g., coronary or cerebrovascular disease
XX or hypertension.
XX Claim 4; SEQ ID NO 1; 37pp; English.
XX The invention relates to a variant dimethylarginine
XX dimethylaminohydrolase-1 (DDAH 1; ADE71464) and a nucleic acid encoding
XX it (ADE71463). DDAHs regulate the metabolism of ADMA (asymmetrical
XX dimethyl-arginine; a naturally occurring nitric oxide synthase inhibitor)
XX and nitric oxide, which are relevant in disorders such as cardiovascular
XX disease and diabetes. The DDAH 1 gene is located on chromosome 1, which
XX has been implicated in susceptibility to diseases such as familial
XX combined hyperlipidaemia, premature coronary artery disease (CAD), non
XX insulin-dependent (type 2) diabetes mellitus, and diastolic hypertension.
XX The variant DDAH 1 of the invention has Met at position 87, rather than
XX Thr, which is caused by a C to T change in exon 1 at a position
XX corresponding to base 260 in ADE71463 and ADE71465. Individuals
XX or homozygous for the Thr87Met allele have an increased risk of
XX developing cardiovascular diseases and diabetes. The invention also
XX relates to a DDAH 1 cDNA capture probe; a method and kit for diagnosing a
XX susceptibility to a cardiovascular disease and diabetes in an individual
XX by determining the DDAH 1 genotype; a method of treating diabetes or
XX vascular complications of diabetes by enhancing nitric oxide
XX availability, production or concentration; a method of targeting
XX treatment of cardiovascular disease and diabetes in a hypertensive

CC patient by determining the DDAAH 1 genotype and treating them with a drug
CC that affects nitric oxide availability, production or metabolism; and a
CC transgenic animal which carries a human variant DDAAH 1 nucleic acid
CC sequence. The nucleic acid molecules and polypeptides are useful for
CC treating cardiovascular disease (e.g., coronary heart disease,
CC cerebrovascular disease, and hypertension), and diabetes and its vascular
CC complications. The methods are useful for determining whether a patient
CC will benefit from treatment with a drug which affects nitric oxide
CC availability, production or metabolism; a drug which reduces ADMA
CC availability or concentration; or an agent which elevates DDAAH
CC availability or concentration (such as DDAAH agonist). The methods are
CC also useful for determining whether a patient will be at risk of adverse
CC effects if DDAAH antagonists are administered. The present sequence
CC represents a specifically claimed nucleic acid encoding the variant
CC (Met87) DDAAH 1.

XX SQ Sequence 858 BP; 218 A; 212 C; 258 G; 170 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6,14e-153 Length: 858
Score: 1433.00 Matches: 282
Percent Similarity: 99.30% Conservative: 1
Best Local Similarity: 98.95% Mismatches: 2
Query Match: 98.56% Indels: 0
DB: 9 Gaps: 0

US-09-889-733B-2 (1-285) x ADE71463 (1-858)

QY 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValAlaCysAla 20
DB 1 ATGGCCGGGGCTGGCCACCGCCGCCCTTCGGCCGGCCACCCACCGCGGTGGTGGCGGCG 60
QY 21 LeuProGluSerLeuGlyHisAlaLeuArgSerAlaLysGlyGluGluValAspVal 40
DB 51 CTACCCGAGTCGCTGGCCAGCAGCAGCGCTGAGAGCGCCAGAGCGGAGGTCGACGTC 120
QY 41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLeuGlyLeu 60
DB 121 GCCCGCGGGAGCAAGCAGCAGCAGCTCTACGTGGCGGTGGTGGCGAGCAAGCTGGGCGTG 180
QY 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
DB 181 CAGGTGGTGGAGCTCCGCCGCCAGCAGAGCCCTCCGAGCTCGCTTCGTGGAGAGAGTG 240
QY 81 AlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLys 100
DB 241 GCCGTGGTGGTGGAGAGATGCCCTCATCACCGACCGCGGGCGCCGAGCGGAGGAAG 300
QY 101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
DB 301 GAGGTGACATGATGAAGAAGCATTAGAAAACCTTCAGCTCAATATAGTAGATGAA 360
QY 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheVal 140
DB 361 GATGAATGCAACTTAGATGGCGAGATGTTTATTTCACAGGCAGAGAGATTTTGTG 420
QY 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp 160
DB 421 GGCCTTTCCAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCTGATCTTTAAGGAC 480
QY 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
DB 481 TATGAGTCTCCACAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 540
QY 181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200
DB 541 GCTGGGCTTAACCTGATCCAAATGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATC 600
QY 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
DB 601 ATGCACACATGAGTACGACCCGCTACGACAAACCTCATCTGTGCTGATGATGATGATGAT 660
QY 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240

DB 661 AACTGTATATCTAATATATCCCAACAAAGGACGCTCTTCTGACACGACCCCGAA 720
QY 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal 260
DB 721 GAGTATCCAGAAAGTGCAGAGGTTTATGAGAACTGAGAGGACCATATGCTGATCCCGTG 780
QY 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuIleAsn 280
DB 781 AGCATGCTGACTGAGAAAGTGGATGGTGTCTCAGCTGCTGCTCAGTCTGCTCAGTTTAA 840
QY 281 LysLysValAspSer 285
DB 841 AAGAAAGTAGACTCC 855

RESULT 8

ADB53014

ID ADB53014 standard; DNA; 3008 BP.

XX ADB53014;

XX ADB53014;

DT 04-DEC-2003 (first entry)

XX 04-DEC-2003 (first entry)

DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3556.

XX toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;

KW toxicity marker; toxicity progression; drug screening;

KW primary rat hepatocyte toxicity modelling; gene; da.

XX Rattus norvegicus.

OS Rattus norvegicus.

XX WO2003065993-A2.

PD 14-AUG-2003.

XX 04-FEB-2003; 2003WO-US003482.

XX 04-FEB-2002; 2002US-0353171P.

PR 13-MAR-2002; 2002US-0363534P.

PR 10-APR-2002; 2002US-0370248P.

PR 10-APR-2002; 2002US-0371134P.

PR 10-APR-2002; 2002US-0371135P.

PR 10-APR-2002; 2002US-0371150P.

PR 11-APR-2002; 2002US-0371413P.

PR 13-APR-2002; 2002US-0373601P.

PR 19-APR-2002; 2002US-0373602P.

PR 22-APR-2002; 2002US-0374139P.

PR 08-MAY-2002; 2002US-0378370P.

PR 09-MAY-2002; 2002US-0378652P.

PR 09-MAY-2002; 2002US-0378653P.

PR 09-MAY-2002; 2002US-0378655P.

PR 09-JUL-2002; 2002US-0394230P.

PR 04-SEP-2002; 2002US-0394253P.

PR 04-SEP-2002; 2002US-0407688P.

PR 28-JAN-2003; 2003US-0442300P.

XX (GENE-) GENE LOGIC INC.

PA Mendrick D, Porter M, Johnson K, Higgs B, Castile A, Orr M;

PI Elashoff M;

XX WPI; 2003-731472/69.

XX Determining if a compound induces a toxic effect on a tissue or cell, for

XX identifying hepatotoxic compounds, comprises comparing a gene expression

XX profile of a tissue or cell sample to a database of Tox mean and non-Tox.

XX mean values.

XX Claim 44; SEQ ID NO 3556; 874pp; English.

XX The present invention describes a method for determining whether a

XX compound induces a toxic effect on a tissue or cell. The method comprises

XX preparing a gene expression profile of a tissue or cell sample exposed to

CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the tox mean and non-tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.

XX
SQ Sequence 3008 BP; 766 A; 842 C; 658 G; 740 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 5,09e-146 Length: 3008
Score: 1379.00 Matches: 268
Percent Similarity: 96.84% Conservatives: 8
Best Local Similarity: 94.04% Mismatches: 9
Query Match: 94.84% Indels: 0
DB: 9 Gaps: 0

US-09-889-733B-2 (1-285) x ADB53014 (1-3008)

QY 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
Db 432 ATGGCGGCGCTCAGCCACCCCTCGTCTTCGGCGGGCCACCCAGCCGCTGGTGGCGGT 491
QY 21 LeuProGluSerLeuGlyHisAlaLeuArgSerAlaGlyGluValValArgVal 40
Db 492 CGCCCGAGTCCCTGTGGCGGCGCTGAGCGCTCCCGGGCGGAGAGGTGATTC 551
QY 41 AlaArgAlaGluArgGlnHisGlnLeuValGlyValGlyValLeuGlySerLeuGlyLeu 60
Db 552 GCTCGCGCTGAGCGCCAGCAGCAGCTCTACGTGGGCGTCTGGCGAGCAAGCTGGGCTG 611
QY 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
Db 612 CAGGTGGTGCAGCTGGCGCGGCGGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 671
QY 81 AlaValValCysGluGluThrAlaLeuLeuThrArgProGlyValProSerArgArgLys 100
Db 672 GCCTGGTGTGGAGAGAGCGCCCTCATCCCGCGCGGGCGCTAGCGCAGGAG 731
QY 101 GluValAspMetValGlyGluAlaLeuGluValGlyLeuGlnLeuAsnLeuValGluMetLys 120
Db 732 GAGGTTCACATGATGAAGAGGCTTTGGAAACATTCAGCTCAACATAGTAGAGTAA 791
QY 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal 140
Db 792 GATGAATATCACTTAGATGGTGGGAGCTCTTATTCAGGCGAGAGTCTTTTGTG 851
QY 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuAlaAspThrPheLysAsp 160
Db 852 GGCTTTCCAAAGGACAAATCAACAGAGTCTGAGATCTTGCTGATCTTCAAGGAC 911
QY 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
Db 912 TAGCGAGTTTCCACAGTCCCGGGCGGATCTTTGCATTTAAAGAGTTCTGACAGT 971
QY 181 AlaGlyProAsnLeuLeuAlaLeuGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200
Db 972 GCTGGCCCCAACCTGATCGCAATAGGTCCAGTGAATCTCGGAGAGGCCCTCAAGATC 1031
QY 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
Db 1032 ATGCAACAGATGATGACCAACCGTATGACAGCTCACTGTATCCGACGACATGGCGGC 1091
QY 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
Db 1092 AACTGTATATATTAAATATATCCCGAAGAGGCGATGCTTGTGTCACCGAACCCAGAA 1151
QY 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal 260

Db 1152 GAGTACCCAGAAAGCGCAAGGTTTATGAGAGCTCAAGGACCATCTACTGATCCCTGTG 1211
QY 261 SerMetSerGluLeuGluLysValAspGlyLeuThrCysCysSerValLeuIleAsn 280
Db 1212 AGCAATCTCAGATGGAAGAGTGGAGCGCTTGCTACCTCGTCTGCTCCGTTTATTAAAC 1271
QY 281 LysLysValAspSer 285
Db 1272 AAGAGACAGACTCT 1286
RESULT 9
ADB79843
ID ADB79843 standard; DNA; 3008 BP.
XX
AC ADB79843;
XX
DT 04-DEC-2003 (first entry)
XX
DE Rat N-G,N-G-dimethylarginine dimethylaminohydrolase DNA, SEQ ID 83.
XX
KW Analgesic; pain; streptozocin-induced diabetes; rat; gene; ds.
XX
OS Rattus norvegicus.
XX
FN EPI279744-A2.
XX
PD 29-JAN-2003.
XX
PF 26-JUL-2002; 2002EP-00255249.
XX
PR 27-JUL-2001; 2001GB-00018354.
PR 07-FEB-2002; 2002GB-00002910.
XX
PA (WARN) WARNER LAMBERT CO.
XX
PI Brooksbank RA, Dixon AK, Lee K, Pinnock RD;
XX
XX WPI; 2003-395407/38.
DR P-PSDB; ADB79842.
XX
PT Use of isolated gene sequences and encoded polypeptides that are
PT upregulated in the spinal cord in response to streptozocin-induced
PT diabetes for screening compounds for the treatment of pain, or for
PT diagnosing pain.
XX
PS Claim 1; Page 155; 334pp; English.
XX
CC The present invention relates to nucleotide sequences which are useful in
CC the screening of compounds for the treatment of pain, or for the
CC diagnosis of pain. The nucleotide sequences are up-regulated in the
CC spinal cord in response to streptozocin-induced diabetes. The present
CC sequence is one such nucleotide sequence.
XX
SQ Sequence 3008 BP; 766 A; 842 C; 658 G; 740 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 5,09e-146 Length: 3008
Score: 1379.00 Matches: 268
Percent Similarity: 96.84% Conservatives: 8
Best Local Similarity: 94.04% Mismatches: 9
Query Match: 94.84% Indels: 0
DB: 9 Gaps: 0

US-09-889-733B-2 (1-285) x ADB79843 (1-3008)

QY 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
Db 432 ATGGCGGCGCTCAGCCACCCCTCGTCTTCGGCGGGCCACCCAGCCGCTGGTGGCGGT 491
QY 21 LeuProGluSerLeuGlyHisAlaLeuArgSerAlaLysGlyGluValValArgVal 40
Db 492 CGCCCGAGTCCCTGTGGCGGCGCTGAGCGCTCCCGGGCGGAGAGGTGATTC 551

```
QY 41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGluVal 60
Db 552 GCTCGCGTGGCGGACAGCAGCTTACGTGGCGTGGCGAGCAGCTGGCGCTG 611
QY 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
Db 612 CAGGTGGTGCAGCTCCGCGCCGACGAGAGCCTGCTGACTGCTGCTGCTGAGAGCAGCTG 671
QY 81 AlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLys 100
Db 672 GCGGTGGTGGAGAGACCGCTCATACCCGCGCGCGCGCTGAGCGGAGAG 731
QY 101 GluValAspMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
Db 732 GAGGTGACATGATGAAAGAGCTTTGAAAACTTCAGCTCAACATAGTAGAGTAA 791
QY 121 AspGluAsnAlaThrLeuAspGlyCysValLeuPheThrGlyArgGluPheVal 140
Db 792 GATGAAATGCAACCTTAGATGGTGGGACGCTCTATTACAGCAGAGAGTTTTTGTG 851
QY 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp 160
Db 852 GGCCTTCCAAAAGACAAATCAACGAGGTGCTGAGATCTTGGCTGATCTTCAAGGAC 911
QY 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
Db 912 TACGAGATTTCCACAGTCCCGTGGCCGATCTTTGCAATTTAAAGAGTTTCTGAGCATG 971
QY 181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200
Db 972 GGTGCCCCACCTGATGCAATAGGTTCAGTGAATCTGCGAGAGGCCCTCAGATC 1031
QY 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
Db 1032 ATGCACACAGATGAGTGACACCGCTTATGCAAGCTCACTGACCGGACGATGCGCGCC 1091
QY 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
Db 1092 AACTGTATATTAAATATCCCGAAGAGGATGCTTGTGACCGACCCAGAA 1151
QY 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal 260
Db 1152 GAGTACCCAGAAAGCGCAAGGTTATGAGAAGCTCAAGGACCATCTACTGATCCCTGTG 1211
QY 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysSerValLeuIleAsn 280
Db 1212 AGCAATCTGAGATGAAAGAGTGAGCGCTGCTCACCTGCTGCTGCTGCTGCTGCTGCTG 1271
QY 281 LysLysValAspSer 285
Db 1272 AAGAAGACAGACTCT 1286
RESULT 10
ADE53981
ID ADE53981 standard; cDNA; 3778 BP.
XX AC ADE53981;
XX DT 29-JAN-2004 (first entry)
XX DE Human prostate cancer cDNA #328.
XX KW Human; prostate cancer; ss; cDNA combination; differential expression;
XX KW gene.
XX OS Homo sapiens.
XX XX US2003190640-Al.
XX PD 09-OCT-2003.
XX XX 29-MAY-2002; 2002US-00252157.
XX PF
```

```
XX 31-MAY-2001; 2001US-0295048P.
XX (FARI/) PARIS M.
XX (PEAR/) PEARSON C I.
XX Paris M, Pearson CI;
XX WPI; 2003-831619/77.
XX New combination comprising cDNAs that are differentially expressed in
XX prostate cancer, useful for diagnosing, treating or monitoring the
XX progression of treatment of prostate cancer.
XX Claim 1; SEQ ID NO 328; 42pp; English.
XX The invention relates to a combination comprising a number of cDNAs
XX expressed in prostate cancer. The invention also relates to a method for
XX detecting differential expression of one or more cDNAs in a sample
XX containing nucleic acids by hybridising a substrate with the nucleic
XX acids, thus forming one or more hybridisation complexes, detecting
XX hybridisation complex formation and comparing the complexes formed with
XX standard complexes, where differences between the standard and the sample
XX complex formation indicate differential expression of cDNAs in the
XX sample. The differential expression is diagnostic of prostate cancer. The
XX invention also relates to proteins and antibodies related to the cDNAs.
XX The combination is useful for diagnosing, treating or monitoring the
XX progression of treatment of prostate cancer. The antibodies are useful
XX for detecting prostate cancer. This sequence represents a human prostate
XX cancer cDNA of the invention.
XX SQ Sequence 3778 BP; 1064 A; 821 C; 719 G; 1150 T; 0 U; 24 Other;
Alignment Scores:
Pred. No.: 1,56e-125 Length: 3778
Score: 1200.00 Matches: 238
Percent Similarity: 99.58% Conservative: 0
Best Local Similarity: 99.58% Mismatches: 0
Query Match: 82.53% Indels: 1
DB: Gaps: 0
US-09-889-733B-2 (1-285) x ADE53981 (1-3778)
QY 48 GlnLeuTyrValGlyValLeuGlySerLysLeuGlyValValGluLeuProAla 67
Db 2 CAGCTTACGTGGCGTGTGGCGCAGCAAGCTGGGGCTGCAGGTGGTGGAGTGGCGGC 61
QY 68 AspGluSerLeuProAspCysValPheValGluAspValAlaValValCysGluIleThr 87
Db 62 GACGAGAGCTTCCGAGCTGCTCTGCTGGAGGACGTGGCGCTGTGTGCGAGGAGAGCG 121
QY 88 AlaLeuIleThrArgProGlyAlaProSerArgLysGluValAspMetLysGlu 107
Db 122 GCGCTCATCCCGACCCCGCGCGCGAGCCGAGAGAGGAGGTTCACATGATGAAGAA 181
QY 108 AlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeuAsp 127
Db 182 GCATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAAGATGAAATGCAACTTTAGAT 241
QY 128 GlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLysArgThrAsn 147
Db 242 GCGGAGATGTTTTATTCACGCGAGAGAAATTTTGTGGGCTTTCCAAAAGGACAAAT 301
QY 148 -GlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrValPr 167
Db 302 CCAACGAGGTGCTGAAATCTTGGCTGATACTTTTAAGGACTATGCAGTCTCCACAGTGCC 361
QY 167 cValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAl 187
Db 362 AGTGCGAGATGGGTGTCATTTGAAGAGTTTTCGAGCATGGTGGGCTTAACTGATGCG 421
QY 187 alleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHi 207
```

422 AATTGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATCATGCAACAGATGAGTGACCA 481
 207 sArgTyrAspLysLeuThrValProAspAspLeuAlaAsnCysIleTyrLeuAsnI 227
 482 CGCTACGACAAACTCACTGCTGCTGATGACATAGCAGCAAACTGTATATATCTAAATAT 541
 227 eProAsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAlaI 247
 542 CCCACAAAGGGCAGCTGCTGCTGCACCGAACCCCGAAGAGATATCCAGAAGTGCAAA 601
 247 sValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGlu 267
 602 GGTATATGAGAACTGAGGACCATATGCTGATCCCGTGAGCATGCTGGAAGTGA 661
 267 sValAspGlyLeuLeuThrCysCysSerValLeuIleAsnLysLysValAspSer 285
 662 GGTGGTGGGCTGCTCACCTGCTGCTCAGTTTATTAACAAGAAAGTAGACTCC 716

RESULT 11

AAAF16077
 ID AAF16077 standard; cDNA; 3651 BP.
 AC AAF16077;
 DT 13-MAR-2001 (first entry)
 DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:512.
 KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
 KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 KW wound; infectious disease; ss.
 CS Homo sapiens.

XX WO200055174-A1.
 XX 21-SEP-2000.
 XX 08-MAR-2000; 2000WO-US005988.
 XX 12-MAR-1999; 99US-0124270P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX (ROSE/) ROSEN C A.
 XX Rosen CA, Ruben SM;
 XX WPI; 2000-587513/55.
 XX P-PSDE; AAB56874.

Prostate cancer associated gene sequences, referred to as prostate cancer
 antigens, useful for treatment, prevention, and diagnosis of disorders
 such as prostate cancer.

Claim 1; Page 976-977; 2338pp; English.

AAAF15566 to AAF16505 encode the human prostate cancer associated
 proteins, called prostate cancer antigens. Given in AAB56363 to AAB57302.
 The prostate cancer antigens can have neuroprotective, cytostatic,
 cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
 nephrotropic, antiinfective, gynaecological and antibacterial activities,
 and can be used in gene therapy. The prostate cancer antigen
 polynucleotides may be used for detection of prostate cancer, diagnosis
 identification, as chromosome markers, and for numerous other chromosomal
 or research purposes. The prostate cancer antigens may be used to treat
 disorders such as neural, immune, muscular, reproductive,
 gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
 AAB57303 represent sequences used in the exemplification of the present
 invention

XX Sequence 3651 BP; 1016 A; 801 C; 684 G; 1141 T; 0 U; 9 Other;
 SQ
 Alignment Scores:
 Pred. No.: 4.24e-110 Length: 3651
 Score: 1064.00 Matches: 212
 Percent Similarity: 99.53% Conservative: 0
 Best Local Similarity: 99.53% Mismatches: 1
 Query Match: 73.18% Indels: 1
 DB: 3 Gaps: 0

US-09-889-733B-2 (1-285) x AAF16077 (1-3651)

QY 73 AspCysValPheValGluAspValAlaValValCysGluGluThrAlaLeuIleThrArg 92
 DB 4 GACTGCGTCTTCGTGGAGGACGTGGCGTGTGTGCGAGGAGAGCGCCCTCATCACCGGA 63
 QY 93 ProGlyAlaProSerArgArgLysGluValAspMetMetLysGluAlaLeuGluLysLeu 112
 DB 64 CCCGGGCGCCGAGCGGAGGAGGTTGACATGATGAAGAAGCATTAGAAAAAATT 123
 QY 113 GlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeuAspGlyGlyAspValLeu 132
 DB 124 CAGCTCAATATAGTAGAGATGAAAGATGAAATCAACTTTAGATGGCGGAGATGTTT 183
 QY 133 PheThrGlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyValaGlu 152
 DB 184 TTCACGSCACAGAAATTTTGTGGCGCTTTCCAAAAGGACAAATCAACGAGGTGCTGAA 243
 QY 153 IleLeuAlaAspThrPheLysAspTyrAlaValSerThrValProValAlaAspGlyLeu 172
 DB 244 ATCTGGCTGTACTATTTTAAGGACTATGTCAGTCTCCACAGTCCAGTGGCAGATGGGTG 303
 QY 173 HisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySerSerGlu 192
 DB 304 CATTTGAGAGTTTCTGCAGCATGGCTGGGCTTAACCTGATCGCAATTGGTCTAGTGAA 363
 QY 193 SerAlaGlnLysAlaLeuLysIleMetGlnMetSerAspHisArgTyrAspLysLeu 212
 DB 364 TCTGCACAGAGGCCCTTAAGATCATGCAACAGATGAGTACCACCGCTACGACAACTC 423
 QY 213 ThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHis 232
 DB 424 ACTGTGCTGTATGATAGATAGAGCAAACTGTATATATCTTAATATCCCCACAAAGGGCAC 483
 QY 233 ValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLysValTyrGluLysLeu 252
 DB 484 GTCTTGCTGCACCGAACCCCGAAGAGTATCCAGAAAGTGCAGAAAGTTTATGAGAAACTG 543
 QY 253 LysAspHisMetLeuIleProValSerMetSerGluLeuGluLysValAspGlyLeuLeu 272
 DB 544 AAGGACCATATGCTGATCCCGTGAGCATGCTGAAGTGAAGAAAGGTGGATGGGCTGCTC 603
 QY 273 ThrCysCysSerValLeuIleAsnLysLysValAspSer 285
 DB 604 ACCTGCTGTCA-GTTTAAATTAACAAGAAAGTAGACTCC 641

RESULT 12

AAAD00677
 ID AAD00677 standard; cDNA; 1228 BP.
 AC AAD00677;
 DT 08-SEP-2000 (first entry)

Human Hydrolase protein-2 (HYDRL-2) encoding cDNA.
 DE
 KW Hydrolase; HYDRL; human; clone 949738; cytostatic; immunosuppressive;
 KW antiinflammatory; neuroprotective; cerebroprotective; anticonvulsant;
 KW nephrotropic; antibody; agonist; antagonist; diagnosis; treatment;
 KW prevention; cell proliferative disorder; cancer; inflammation; AIDS;
 KW Acquired Immune Deficiency Syndrome; autoimmune/inflammatory disorder;
 KW neurological disorder; epilepsy; stroke; medullary sponge kidney;

renal disorder; adrenal disorder; adrenoleucodystrophy; ss.

Homo sapiens.

Location/Qualifiers

263..1120

/tag= a

/product= "Human Hydrolase protein-2"

/note= "Derived from PANCN0105 library"

323..373

/tag= b

/note= "DNA used in hybridisation and amplification

technology for identification of HYDRL sequences"

1067..1117

/tag= c

/note= "DNA used in hybridisation and amplification

technology for identification of HYDRL sequences"

WO200028045-A2.

18-MAY-2000.

12-NOV-1999; 99WO-US027009.

12-NOV-1998; 98US-0172256P.

21-MAY-1999; 99US-0135519P.

(INCY-) INCYTE PHARM INC.

Tang TY, Hillman JL, Yue H, Lal P, Bandman O, Corley NC;

Guegler KJ, Baughn MR, Lu DAM, Azimzai Y, Yang J;

WPI; 2000-376557/32.

P-PSDB; AAY71104.

Novel human hydrolase protein useful for diagnosing, treating and preventing cell proliferative, autoimmune and inflammatory, neurological, renal, adrenal and genetic disorders.

Claim 9; Page 94-95; 106pp; English.

The present sequence is the human hydrolase protein-2 (HYDRL-2) encoding CDNA, identified in Inocyte clone 949738, derived from the PANCN0105 library. It is expressed in reproductive, gastrointestinal, nervous, and cardiovascular tissues. This sequence maps to chromosome 6 within the interval from 42.30 to 45.40 centiMorgans (cM), to chromosome 9 within the interval from 130.40 to 166.50 cM, and to chromosome 16 within the interval from 88.10 to 92.60 cM. HYDRL antibodies, agonists and antagonists are useful for diagnosis, treatment and prevention of disorders associated with altered expression or activity of HYDRL. It includes cell proliferative disorder such as cancer, autoimmune or inflammatory disorders such as inflammation, AIDS, neurological disorder such as epilepsy, stroke, renal disorder such as medullary sponge kidney, adrenal disorder such as adrenoleucodystrophy. The nucleotide sequences are useful for detecting and quantifying gene expression in tissues

Sequence 1228 BP; 237 A; 359 C; 418 G; 214 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 9,996-69 Length: 1228

Score: 694.00 Matches: 146

Percent Similarity: 68.93% Conservative: 47

Best Local Similarity: 52.14% Mismatches: 83

Query Match: 47.73% Indels: 4

DB: 3 Gaps: 3

US-09-889-733B-2 (1-285) x RAD00677 (1-1228)

3 GlyLeuGlyHisPro--SerAlaPheGlyArgAlaThrHisAlaValValArgAlaLeu 21

260 CGGATGGGACCGCGGGGAGGGCTGGCGCTGCTCCCATGCTCCGATGCGGGGAGTC 319

22 ProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluGluValAspValala 41

320 CCAGAGACCTGGCGTCCGGGGAAGGTCCGGGGCTGCCTCCGCTCTGATCTGGCC 379
42 ArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGln 61
380 AAAGCTCAAGGGGACACGGGTGCTGGGAGGTAAACTGAGGACACGACTGGGGGTACAG 439
62 ValValGluLeuProAlaAspLysSerLeuProAspCysValPheValGluAspValala 81
440 CTGCTAGAACCTGCCACCTGAGGAGTCAATGGCGTGGGACCGCTGCTGGGACACGACC 499
82 ValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLysGlu 101
500 GTGATCCAAAGGGGACACGGCCCTAATCACGGGGCTCCGAGCCCGCTCGTAGGCGACAG 559
102 ValAspMetMetLysGluAlaLeuLysLeuGlnLeuAsnLeuValGluMetLysAsp 121
560 GTGATGAGGATCCGCAAGCCCTGACAGACCTGGGGCTCCGAATTTGTGGAATAGGACAC 619
122 GluAsnAlaThrLeuAspGlyValLeuPheThrGlyArgGluPhePheValGly 141
620 GAGAACGGGACGCTGGATGGCACTGACGTTCTTTCACCGGCGGGAGTTCCTGATGCG 679
142 LeuSerLysArgThrAsnGlnArgGlyAlaGluLeuLeuAlaAspThrPheLysAspTyr 161
680 CTCTCCAAATGGACCAATCACCGAGGAGCTGAGATCGTGGGGGACACGTTCCGGGACTTC 739
162 AlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetala 181
740 GCGGTCTCCAGTGTCCAGTCTCGGGTCCCTCCACCTGGCGGTCTCTGGGCGATGGG 799
182 GlyProAsnLeuLeuAlaLeuGlySerSerGluSerAlaGlnLysAlaLeuLysLeuMet 201
800 GGACCTCGCACTGTTGTGGCGGACGACGCGTCCCAAGGCTGTCCGGGCAATG 859
202 GlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsn 221
860 GCAGTGTGACAGATACCCATATGCTCCCTGACCTCCACGATGACGAGCTCTGAC 919
222 CysIleTyrLeu-----AsnIleProAsnLysGlyHisValLeuLeuHisArgThrPro 239
920 TGTCTCTCTCTGCTGCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 979
240 GluGluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuLeuPro 259
980 GGGGATCTGCCCAACAGCAGCAGGAGGACACTGCGAGAAGCTCTCTGATGTACCTGTACT 1039
260 ValSerMetSerGluLeuGluLysValAsp---GlyLeuLeuThrCysCysSerValLeu 278
1040 GTGCTCTGCTCAGAACTGGAGAAGCGCGGCGCGGCTCAGCTCCCTCTGCTGTGTGCTC 1099
RESULT 13
AAZ52967
ID AAZ52967 standard; cDNA; 1376 BP.
XX
AC AAZ52967;
XX
DT 14-MAR-2000 (first entry)
XX
DE Human prostate tumor cDNA library derived EST fragment #110.
XX
KW Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
XX treatment; ds.
XX
OS Homo sapiens.
XX
FN DE19820190-A1.
XX
PD 04-NOV-1999.
XX
PF 28-APR-1998; 98DE-01020190.
XX
PR 28-APR-1998; 98DE-01020190.

[illegible]

QY	202	GlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsn	221
Db	894	CGAGTCTGTCAGAGATCACCATATGCTCTGCTGACCTCCAGATGACGAGTGTGTGAC	953
QY	222	CysIleTyrLeu-----AsnIleProAsnLysGlyHisValLeuLeuHisArgThrPro	239
Db	954	TGTCCTCTTCTTCGTCCTGGGTTCCTGCTGTGTGCCCTTTTCCTCTGCACCTGGAGGT	1013
QY	240	GlugLysTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuLeuPro	259
Db	1014	GGGGAATCTGCCAACAGCAGGAGGCACTGCAGAACTCTCTGATGTCACCTTGGTACCT	1073
QY	260	ValSerMetSerGluLeuGluLysValAsp--GlyLeuLeuThrCysCysSerValLeu	278
Db	1074	GTCTCCTGCTCAGAACTGGAAGAGGTGGCGCGGGCTCAGCTCCCTCTGCTTGGTGCTC	1133
RESULT 14			
AAA47656			
ID	AAA47656 standard; cDNA; 858 BP.		
XX			
XX	AAA47656;		
XX			
XX	08-NOV-2000 (first entry)		
DE	Dimethylarginine dimethylaminohydrolase (DDAH2) coding sequence.		
XX			
KW	Dimethylarginine dimethylaminohydrolase; DDAH, DDAH1, DDAH2;		
KW	arginine deaminase; hyperlipidemia; renal failure; hypertension;		
KW	restenosis; atherosclerosis; schizophrenia; multiple sclerosis; cancer;		
KW	ischemia reperfusion injury; septic shock; multi organ failure;		
KW	arthritis; skin disorders; inflammatory cardiac disease; migraine;		
KW	infection; ds.		
OS	Homo sapiens.		
XX			
XX			
FH	Key Location/Qualifiers		
FT	CDS 1..858		
FT	/*tag= a		
FT	/product= "DDAH2"		
XX			
PN	WO200004888-A2.		
XX			
PD	03-AUG-2000.		
XX			
XX	26-JAN-2000; 2000WO-GB000226.		
XX			
PR	26-JAN-1999; 99GB-00001705.		
PR	04-JUN-1999; 99GB-00013066.		
XX			
PA	(UNLO) UNIV COLLEGE LONDON.		
XX			
FI	Vallance PJT, Leiper JM, Whitley GSJ, Charles IG;		
XX			
XX	WPI; 2000-543392/49.		
DR	P-PSDB; AAB01478.		
XX			
PT	Novel methylarginase polypeptides and polynucleotides, used to identify		
PT	modulators of them, which are used in the treatment of e.g. cancer,		
PT	hypertension, and bacterial infections.		
XX			
PS	Claim 1; Page 57-58; 68pp; English.		
XX			
CC	Nucleotides encoding methylarginase polypeptides, vectors comprising		
CC	these nucleotides and the polypeptides themselves can be used in		
CC	medications for the treatment of hyperlipidemia, renal failure,		
CC	hypertension, restenosis after angioplasty, atherosclerosis,		
CC	complications of heart failure, schizophrenia, multiple sclerosis or		
CC	cancer. Modulators of the enzyme can be used in medicaments for the		
CC	treatment of ischemia-reperfusion injury of the brain or heart, cancer,		
CC	lethal hypertension in severe inflammatory conditions such as septic		
CC	shock or multi-organ failure, or local and systemic inflammatory		
CC	disorders including arthritis, skin disorders, inflammatory cardiac		
CC	disease, migraine, or microbial or bacterial infection		

XX SQ Sequence 858 BP; 150 A; 268 C; 281 G; 159 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,87e-68 Length: 858
Score: 688.00 Matches: 145
Percent Similarity: 68.82% Conservative: 47
Best Local Similarity: 51.97% Mismatches: 83
Query Match: 47.32% Indels: 4
DB: 3 Gaps: 3

US-09-889-733B-2 (1-285) x AAA47656 (1-858)

QY 4 LeuGlyHisPro---SerAlaPheGlyArgAlaThrHisAlaValValArgAlaLeuPro 22
DB 1 ATGGGACGCGGGGAGGGGCTGGCGCTGCTCCATCCCTGATCGGGGAGTCCCA 60
QY 23 GluSerLeuCysGlnHisAlaLeuArgSerAlaGlyGluGluValAlaArg 42
DB 61 GAGAGCTGCGCTGGGGAGGAGTGGGGCTGGCGCTCCCTCTGATCTGGCCAA 120
QY 43 AlaGluArgGlnHisGlnLeuTyValGlyValLeuGlySerLysLeuGlyLeuGlnVal 62
DB 121 GCTCAAGGGAGCAGCGGGTGGGAGGTAACTGAGGCAACGACTGGGGCTACAGCTG 180
QY 63 ValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAlaVal 82
DB 181 CTGAAGTGCACCTGAGGAGTCAATTCCTGCGGACCGCTGCTGGCGACACGCGCTG 240
QY 83 ValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgGlyGluVal 102
DB 241 ATCAAGGGAGCAGCGGCTTAATCACCGGCTTGGAGCCCTGAGCCCTGTAGCCAGAGTTC 300
QY 103 AspMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGlu 122
DB 301 GATGAGTCCGCAAGCCCTGCAAGACCTGGGCTCCGAATTTGGAAATAGGAGACGAG 360
QY 123 AsnAlaThrLeuAspGlyValValLeuPheThrClyArgGluPhePheValGlyLeu 142
DB 361 AACCGAGCTGTGGAGGAGTCACTGAGTCTCTTACCGCGCGGAGTTCCTGAGGCTTC 420
QY 143 SerLysArgThrAenGlnArgGlyAlaGluLeuAlaAspThrPheLysAspTyAla 162
DB 421 TCCAAATGGACCAATCACCGAGGAGTGAATGTCGGCGGACAGTTCGGGACTTCGCG 480
QY 163 ValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGly 192
DB 481 GTCTCCACTGTGCGAGTCTCGGCTCCCTCCACCTGCGCGGTCTCTCGCGCATGGGGGA 540
QY 183 ProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGln 202
DB 541 CCTGCGACTGTGTGGCAGGAGCAGCAGCGCTGCCCAAGAGCTGTCCGGCAATGGCA 600
QY 203 GlnMetSerAspHisArgTyAspLysLeuThrValProAspAspIleAlaAlaAsnCys 222
DB 601 GTGCTGACAGATCACCATATGCTCCCTGACCTCCACCTCCAGATGACGAGCTGTGACTGT 660
QY 223 IleTyLeu-----AsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
DB 661 CTCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 241 GluTyProGluSerAlaLysValTyGluLysLeuLysAspHisMetLeuIleProVal 260
DB 721 GATGCTCCCAACGCGAGGAGGAGTCTGATGATGATGATGATGATGATGATGATGATGATG 780
QY 261 SerMetSerGluLeuGluLysValAsp---GlyLeuLeuThrCysCysSerValLeu 278
DB 781 TCCGTCTCAGAACTGGAGAAAGCTGGCGCGGCTCAGCTCCCTCTCTCTGCTGCTC 837

RESULT 15

AAA99390

ID AAA99390 standard; cDNA; 845 BP.

XX

AC AAA99390;
XX 22-JAN-2001 (first entry)
DT Human melanoma growth related factor-2 cDNA sequence.
XX Melanoma growth related factor-2; MGRF-2; human; ss.
XX Homo sapiens.
OS CN1257921-A.
PN 28-JUN-2000.
XX 18-DEC-1998; 98CN-00125527.
PF 18-DEC-1998; 98CN-00125527.
PR (UYFU-) UNIV FUDAN.
XX Yu L, Fu Q, Zhang H;
PI WPI; 2000-544294/50.
XX P-PSDB; AAB26802.
XX Human melanoma growth correlation factor and its coding sequence,
PT preparing process and usage.
XX Claim 1; Page 14; 20pp; Chinese.
XX The present invention relates to a melanoma growth related factor (MGRF-
CC 2) cDNA and protein sequence, isolated from humans. The present sequence
CC represents the cDNA of the invention
XX Sequence 845 BP; 152 A; 250 C; 285 G; 158 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,56e-66 Length: 845
Score: 669.50 Matches: 141
Percent Similarity: 59.52% Conservative: 46
Best Local Similarity: 52.42% Mismatches: 78
Query Match: 46.05% Indels: 4
DB: 3 Gaps: 2

US-09-889-733B-2 (1-285) x AAA99390 (1-845)

QY 3 GlyLeuGlyHisPro---SerAlaPheGlyArgAlaThrHisAlaValValArgAlaLeu 21
DB 41 GGGATGGGAGCGCGGGGAGGGGCTGGCGCTGCCATGCCCTGATCCGGGAGTC 100
QY 22 ProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluValAlaAspValAla 41
DB 101 CCAGAGAGCTGGCTGCGGGGAGGTGCGGGGCTGGCGCTTCCGCTCTGATCTGGCC 160
QY 42 ArgAlaGluArgGlnHisGlnLeuTyValGlyValLeuGlySerLysLeuGlyLeuGln 61
DB 161 AAAGCTCAAGGGAGCAGCGGGTGTGGGAGGTAAACTGAGGCAACGACTGGGGCTACAG 220
QY 62 ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAla 81
DB 221 CTGCTAGACTGCCACTTGAGGAGTCAATGCGCTGGGACCGCTGCTGGCGACACGAGC 280
QY 82 ValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGlu 101
DB 281 GTGATCCAGGGGACACGCGCTTATCACGGGCGCTGGAGCGCGCTCTGAGGCCAGAG 340
QY 102 ValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAsp 121
DB 341 GTGATGAGTCCGCAAGCCCTGCAAGACTGGGGCTCCGATTTGTGGAATAGGAGAC 400
QY 122 GluAsnAlaThrLeuAspGlyValValLeuPheThrGlyArgGluPhePheValGly 141
DB 401 GAGAAGCGAGCGCTGATGGCACTGACGTTCTCTTACCGCGCGGAGTTTTTTCGTAGCC 460

QY	142	LeuSerIysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyr	161
Db	461	CTCTCCAAATGGACCAATCACCAGAGAGCTGAGATCGTGGGACACGTTCCGGACTTC	520
QY	162	AlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAla	181
Db	521	CCCGTCTCCACTGTGCAGTCTCGGTCCTCCACCTGCGGGTCTCTGCGGCATGGGG	580
QY	182	GlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMet	201
Db	581	GGACCTCGCACTGTTGTGGCAGCAGCAGCGCTGCCCAAAGGCTGTCCGGGCAATG	640
QY	202	GlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsn	221
Db	641	GCAGTCTGACAGATCACCATATGCTCTCCCTGACCCCTCCAGATGACCGACTGCTGAC	700
QY	222	CysIleTyrIleu-----AsnIleProAsnLysGlyHisValLeuLeuHisArgThrPro	239
Db	701	TGTCTCTTTCTTCTGCTGGGTTCCTGTGTGCCCTTTCTCTCTGACCGTGGAGGT	760
QY	240	GluGluTyrProGluSerAlaLysVal-TyrGluLysLeuLysAspHisMetLeuIlePr	259
Db	761	GGGATCTGCCCAACAGCCAGGAGGACCTGCAGAAAGCTCTGTATGTACCCCTGTTACC	820
QY	259	oValSerMetSerGluLeuGluLys	267
Db	821	TGTGCTCTGCTCAGAACTGGAGAA	845

Search completed: June 7, 2004, 23:27:29
Job time : 387 secs

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 7, 2004, 23:12:01 ; Search time 2529 Seconds
(without alignments)
3365.252 Million cell updates/sec

Title: US-09-889-733B-2
Perfect score: 1454
Sequence: 1 MAGLHPSAFGRATHAVRA.....EKVGLLTCCSLVINKKVD 285

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=slp
-MODE=frame+ p2n.model -DEV=slp
-Q=/cgn2_1/USPTO_spool_p/US09889733/runat_07062004_083528_7982/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=act -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcr -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09889733 @CGN_1_1_5180@runat_07062004_083528_7982 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_estc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_pbg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1389	95.5	1109	11	AK010430	Mus muscu
2	1217	83.7	1001	13	BY709864	BY709864
3	1127	77.5	746	14	CB249709	UI-M-EXD-
4	1125	77.4	862	13	BQ899468	AGENCOURT
5	1118	76.9	889	12	BI763990	603049879
6	1077	74.1	760	14	CA512485	CA512485
7	1064	73.2	822	9	AI647789	UI-R-FUO-
8	1056	72.6	656	13	BY742062	uk43505.x
9	975.5	67.1	860	14	CF548604	AGENCOURT
10	956.5	65.8	800	9	AA986117	uc81h06.x
11	945	65.0	638	13	BY754930	BY754930
12	925	63.6	776	14	CD751679	CD751679
13	896	61.6	731	10	BB611533	BB611533
14	892.5	61.4	897	14	CK019367	AGENCOURT
15	892	61.3	923	12	BI332932	602982027
16	881	60.6	522	9	AI751463	cn10b04.x
17	870	59.8	732	14	CD750895	CD750895
18	867.5	59.7	777	14	CD349526	UI-M-PYO-
19	864	59.4	645	29	CG477727	CG477727
20	857	58.9	685	14	CB247154	UI-M-FIO-
21	853.5	58.7	802	14	CA588435	CA588435
22	822	56.5	501	29	CG617513	OST311187
23	820	56.4	702	13	BY731938	BY731938
24	816.5	56.2	1035	10	BF579996	BF579996
25	806.5	55.5	788	14	CB248156	CB248156
26	796	54.7	825	14	CD752232	CD752232
27	794.5	54.6	654	9	AU296050	AU296050
28	789	54.3	722	14	CK142177	CK142177
29	785	54.0	486	9	AI463821	va31a06.y
30	772	53.1	455	29	AY406329	AY406329
31	771	53.0	455	29	AY406330	Pan trogl
32	770.5	53.0	806	9	AU051351	AU051351
33	764	52.5	617	12	BI221174	BI221174
34	756	52.0	681	12	BM963225	BM963225
35	746	51.3	569	9	AI265716	uj05e03.x
36	744	51.2	556	12	BM964027	BM964027
37	741	51.0	455	29	AY406331	Mus muscu
38	718	49.4	432	10	BE864630	UI-M-BH1-
39	711.5	48.9	867	14	CD363134	CD363134
40	708.5	48.7	783	14	CK138561	CK138561
41	706.5	48.6	639	13	BQ399610	BQ399610
42	705	48.5	547	9	AI661738	va31a06.x
43	701	48.2	624	10	AW514143	hg24a02.x
44	697.5	48.0	776	14	CB317151	CB317151
45	694	47.7	1351	11	AF070667	Homo sapi

ALIGNMENTS

RESULT 1
AK010430
LOCUS
DEFINITION
AK010430
Mus musculus ES cells cDNA, RIKEN full-length enriched library,
clone:241006N07 product:dimethylarginine dimethylaminohydrolase 1,
full insert sequence.
1109 bp mRNA linear HTC 20-SEP-2003
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AK010430.1 GI:12845867
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata;

Db 194 CGCGCCGAGTCCCTGTGCGCCACCGCTGAGCGCTCGAGGGCGAGAGGTGGATTTC 253
 Qy 41 AlaArgAlaGluArgGlnHis-GlnLeuTyr-ValGlyValLeuGlySerLysLeuGlyL 60
 Db 254 GCTCGCGCGAGCGCCAGCACCGACTCTACGCTGGCGGTCTGGCGAGCAAGCTGGGCG 313
 Qy 60 euGlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspV 80
 Db 314 TCAGAGTGTGCGAGTGCCTGCGCGCGAGAGCGCTGCCGAGCTGCGTGTCTGAGGAGCG 373
 Qy 80 alAlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgL 100
 Db 374 TGGCGCTGTGTGCGAGAGAGCGCCCTCATCCCGCGCGCGCGCGCGCGAGGA 433
 Qy 100 ysGluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnLleValGluMetL 120
 Db 434 AGGAGGTTTCACATGATGAAGAGGCTTTGGAAACCTTCAGCTCAACATAGTAGAGATGA 493
 Qy 120 ysAspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheV 140
 Db 494 AAGATGAAATGCAACTTTGGATGTGGGACGCTCTATTTCAGGCGCAGAGAAATTTTGG 553
 Qy 140 alGlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuLeuAlaAspThrPheLysA 160
 Db 554 TGGCGCTTTCCAAAAGAACAAATCAACAGAGTGTCTGAAATCTTGGCTGATCTTTAAGG 613
 Qy 160 spTyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerM 180
 Db 614 ACTAGCGAGTCTCTACAGTCCCTGCGCGGATCTTTGCATTTAAGAGTTTCTGCGACA 673
 Qy 180 elAlaGlyProAsnLeuLeuAlaLleGlySerSerGluSerAlaGlnLysAlaLeuLysI 200
 Db 674 TGGCGGACCCCAACCTGATTGCAATAGGTCGAGGAACTCTGCAGAGGCGCTCAAGA 733
 Qy 200 leMetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspLleAlaA 220
 Db 734 TCATGCAACAGATGATGACCATCTTTATGCAAGCTCACTGTACCCGACACATGCGCG 793
 Qy 220 laAsnCysLleTyrLeuAsnLleProAlaLysGlyHisValLeuLeuHisArgThrProG 240
 Db 794 CCAACTGCATATATCTAAATATCCCGACGATAGGCGCATGCTTGTGCTGACGTGAACCC 853
 Qy 240 luGluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuLleProv 260
 Db 854 AAGAGTACCCAGAAAGCGCAACGCTCTATGAGNAACCTCAAGGACCATCTACTGATCTGT 913
 Qy 260 alSerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuLle 279
 Db 914 TGAGCAC-TCGGAGATGGAAGTAGCGAGCGCC---GTCCCTGTGCGC-TCCGTTTATTAT 967

 RESULT 3
 CB249709
 LOCUS
 DEFINITION
 UI-N-EX0-bv1-1-07-0-UI.r1 NIH_BMAP_EX0 Mus musculus CDNA clone
 IMAGE:5719830 5', mRNA sequence.
 CB249709
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 746)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs@mail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: pYX-5
 Location/Qualifiers
 1. 746
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="TMAGE:5719830"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (11 phage resistant)"
 /clone_lib="NIH BMAP_EX0"
 /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaïdo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA
 tail, is GTCGCTGGGA. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP). 'Gene discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
 Pred. No.: 4,88e-122 Length: 746
 Score: 1127.00 Matches: 219
 Percent Similarity: 97.83% Conservative: 6
 Best Local Similarity: 95.22% Mismatches: 5
 Query Match: 77.51% Indels: 0
 DB: 14 Gaps: 0

 US-09-889-733B-2 (1-285) x CB249709 (1-746)
 Qy 56 SerLysLeuGlyLeuGlnValGluLeuProAlaAspGluSerLeuProAspCysVal 75
 Db 3 AGCAAGCTGGGCTGCGAGTGTGCGAGCTGCCCGCGAGAGCGCTGCCAGCTGCCGTG 62
 Qy 76 PheValGluAspValAlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAla 95
 Db 63 TTCGTGGAGACGTGGCGCTGTGTGCGAGGAGAGCGCCCTCATCCCGCGCGCGGGCG 122
 Qy 96 ProSerArgArgLysGluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsn 115
 Db 123 CCCAGCGCGAGGAGGAGGTTCACATGATGAAAGAGGCTTTGAAAAAATTCAGCTCAAC 182
 Qy 116 IleValGluMetLysAspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGly 135
 Db 183 ATAGTAGATGAAGATGAAATGCACTTTGGATGGTGGGAGCGCTCTATTTCACAGCG 242
 Qy 136 ArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuLeuAla 155
 Db 243 AGAGATTTTGTGGCGCTTTCCAAAAGACAAATCAACGAGGTGCTGAAATCTTGCT 302
 Qy 156 AspThrPheLysAspTyrAlaValSerThrValProValAlaAspClyLeuHisLeuLys 175
 Db 303 GATPACTTTTAAGACTACCGAGTCTCTACAGTCCCTGTGGCGGATTTCTTTGATTTAAG 362
 Qy 176 SerPheCysSerMetAlaGlyProAsnLeuLeuAlaLleGlySerSerGluSerAlaGln 195
 Db 363 AGTTTCTGAGCATGCGCGGACCCCACTGATTGCAATAGGTTCCAGCGAATCTGCACAG 422
 Qy 196 LysAlaLeuLysLleMetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValPro 215

```

Db      423  AAGGCCCTCAAGATCATGCAACAGATGATGATGACCATCGTTATGACAAGCTCACTGATCCC 482
QY      216  AspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeu 235
Db      483  GACGACATGCGCGCCCACTGCATATATCTAATATATCCCGAGCANAGGCGATGCTTGTG 542
QY      236  HisArgThrProGluGluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHis 255
Db      543  CACCGAACCCCGAAGAGTACCCAGAAACGCGANAGGCTCTATGAGAAACTCAAGACCAT 602
QY      256  MetLeuIleProValSerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCys 275
Db      603  CTAAGTATCCCTGTGAGCAACTCGGAGATGGAAGTGGAGCGGCTTGTCTACCTGCTGC 662
QY      276  SerValLeuIleAsnLysLysValAspSer 285
Db      663  TCGTGTATTTTAAACAAGAAGATAGACTCC 692

RESULT 4
BQ899468
LOCUS      BQ899468      862 bp      mRNA      linear      EST 16-AUG-2002
DEFINITION AGENCOURT_8750357 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6332815
5', mRNA sequence.
ACCESSION BQ899468
VERSION   BQ899468.1 GI:22291482
KEYWORDS  ZST.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
REFERENCE 1 (bases 1 to 862)
AUTHORS  NIH-MGC http://mgi.nci.nih.gov/
JOURNAL  National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.
cDNA Library Preparation: ResGen, Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13791 row: a column: 08
High quality sequence stop: 704.

FEATURES
source
Location/Qualifiers
1..862
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6332815"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_130"
/note="Organ: oocytes; Vector: pCMV-Sport6.1; Site: 1: EcoRV; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 1,07e-121 Length: 862
Score: 1125.00 Matches: 226
Percent Similarity: 96.72% Conservative: 10
Best Local Similarity: 92.62% Mismatches: 5
Query Match: 77.37% Indels: 3
DB: 13 Gaps: 0

US-09-889-733B-2 (1-285) x BQ899468 (1-862)
QY      1  MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValAlaGala 20
Db      131  ATGCCGGGCTCGGCCACCCCTTCGGCTTCGGCGCGGCCACCCACCGCTGCTGGGGCT 190

```

```

QY      21  LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluValValAspVal 40
Db      191  CCGCCCGAGTCTGTGTGCGCCACCGCTGAGGCGCTCGCAGGCGGAGGAGTGGATTTC 250
QY      41  AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
Db      251  GCTCGCCCGAGCGCCAGCAGAGCTCTAGTGGCGCTGCTGGSCACAGCTGGGGCTG 310
QY      61  GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
Db      311  CAGGTGTGTGACGTGCTCCCGCGAGAGAGCTGCGGACTGCGTGTTCGTGGAGCAGGTG 370
QY      81  AlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
Db      371  GCCGTGCTGTGCGAGAGAGCGCCCTCATCCCGCGCGCGCGCGCGCGCGCGAGG 430
QY      101  GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
Db      431  GAGGTTCACATGATGAAGAGGCTTTGGAAAACTTCAGCTCAACATAGTAGAGATGAAA 490
QY      121  AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheVal 140
Db      491  GATGAAAAATGCACTTTGGATGCTGGGAGCTCTATTTCACAGCAGAGAGATTTTTTGTG 550
QY      141  GlyLeuSerLysArgThrAsnGlnArgGlyValAlaGluLeuLeuAlaAspThrPheLysAsp 160
Db      551  GGCCTTTCCAAAAAGAACAAATCAACAGAGGCTGCTGAAATCTTGGCTGATCTTTAAGGAC 610
QY      161  TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
Db      611  TACGCAGTCTCTACAGTCCCTGTGGCGGATCTTTTGCATTAAAGAGTCTTCTGCAGCATG 670
QY      181  AlaGlyProAsnLeuIleAlaIleGlySerSerSerAlaGlnLysAlaLeuLysLeu 200
Db      671  GCCGAGCCCAACCTGATTCGAATAGGCTCCAGCAATCTGCACAGAAGGCCCTCAAGATC 730
QY      201  MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
Db      731  ATGCAACAGATGATGATGACCATCGTTATGACAAGCTCACTGTACCCGAGCATGCGCGCC 790
QY      221  AsnCysIleTyrLeuAsnIle-ProAsnLysGlyHis-ValLeu-LeuHisArgThrPro 239
Db      791  CACTGCATATATCTAAATATATATATATATATATATATATATATATATATATATATAT 850
QY      240  GluGlu 241
Db      851  CAAAAG 856

RESULT 5
BI763990
LOCUS      BI763990      889 bp      mRNA      linear      EST 25-SEP-2001
DEFINITION 603049879F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5189970 5', mRNA sequence.
ACCESSION BI763990
VERSION   BI763990.1 GI:15755568
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 889)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

```


http://image.llnl.gov
 Plate: LUAM11474 row: n column: 19
 High quality sequence start: 3
 High quality sequence stop: 847.

FEATURES

Location/Qualifiers

source

1..889

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5189970"
 /lab_host="DH10B"
 /clone_lib="NHG_MGC 116"
 /note="Organ: pooled colon, kidney, stomach; Vector: PCMV-SPORE5; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred No.: 7,53e-121 Length: 889
 Score: 1118.00 Matches: 237
 Percent Similarity: 92.45% Conservative: 8
 Best Local Similarity: 89.43% Mismatches: 14
 Query Match: 76.89% Indels: 6
 DB: 12 Gaps: 2

US-09-889-733B-2 (1-285) x B1763990 (1-889)

QY 1 MetAlaGlyLeuGlyHisProSerAlaPheClyArgAlaThrHisAlaValValArgAla 20
 DB 102 ATGGCCGGGCTCGGCACCGCCCGCCCTTCGGCCGGCCACCCACCGCGTGGTGGCGGCG 161
 QY 21 LeuProGluSerLeuGlyHisAlaLeuArgSerAlaLysGlyGluValAlaVal 40
 DB 162 CTACCCGAGTCGTCGGCAGCAGCGCTGAGAGCCGAGGCGGAGGAGGTGACGTC 221
 QY 41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
 DB 222 GCGCCGCGGGAACGCGCAGCAGCTACGTCGGCGCTGCTGGCAGCAAGCTGGGCGCTG 281
 QY 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
 DB 282 CAGGTGGTGGAGCTCCCGCCGACGAGAGCTTCGGACTCGCTTCGTGGAGGACGTG 341
 QY 81 AlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLys 100
 DB 342 GCCGTGGTGGAGGAGAGCGCCCTCATCCCGACCGCGGCGCGGAGCGGAGGAG 401
 QY 101 GluValAspMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
 DB 402 GAGGTGATCATGTAAGAAGACATTAAGAACTTCAGCTCAATATAGTAGATGAAA 461
 QY 121 AspGluAsnAlaThrLeuAspGlyValLeuPheThrGlyArgGluPhePheVal 140
 DB 462 GATGAATGCAACTTATAGTGGCGAGATGTTTATTCACAGGACAGATTTTGTG 521
 QY 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp 160
 DB 522 GCGCTTTCCAAAGGACAAATCAACGAGGTGCTGAAATCTTTGGTGTACTTTTAAAGAC 581
 QY 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSer-PheCysSerMe 180
 DB 582 TATGAGTCTCCACAGTCCAGTGCAGATGGGTGGATTTTGAAGAGTTTCTGCACAT 641
 QY 180 TalacGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200

DB 642 GGCTGGGCTAACCTGATCGCAATGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGAT 701
 QY 200 eMetGlnGln-MetSerAspHisArgTyrAsp-LysLeuThrVal-ProAspAspIleAla 219
 DB 702 CATGCAACAGATTGATGACCAAGCGGTACGACAAACTCACTGTGGCTGATGACCTTAG 761
 QY 219 aAlaAsnCysIleTyr--LeuAsnIle--ProAsnLysGlyHisValLeuLeuHisArg 237
 DB 762 CAGCAAAAGTGTATATATCTTAAATATATCCCAACAAGGCGCGTCTTCTTGGAGC 821
 QY 237 gThrProGluGluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetIle 257
 DB 822 GAACCCGGAAGAGTATCCAGAAAGTGCAAGGTTATGAACAACTTGAGGACCTAATGCT 881
 QY 257 urlePro 259
 DB 882 GATCCCC 888
 RESULT 6
 LOCUS CAS12485
 DEFINITION UI-R-FJ0-cpz-1-06-0-UI-r1 UI-R-FJ0 Rattus norvegicus cDNA clone
 ACCESSION CAS12485
 VERSION CAS12485.1 GI:25003439
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 760)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 REVERSE.
 Location/Qualifiers
 1..760
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-FJ0-cpz-1-06-0-UI"
 /tissue_type="embryo"
 /dev_stage="embryo"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-R-FJ0"
 /note="Vector: pX-Asc; Site 1: EcoR I; Site 2: Not I;
 UI-R-FJ0 is a cDNA library containing the following
 tissue(s): rat embryo. The library was constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pT73-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is

644 ACTTACGCTCAACATAGTACAGATGAGATGAAATGCAACTTTGGATGGTGGGACGT 585
 131 lLeuPheThrGlyArgGluPheValGlyLeuSerLysArgThrAsnGlnArgGlyAl 151
 584 CCTATTACAGGAGAGATTTTGTGGCCCTTCCAAAGAACAAATCAACAGAGGTGC 525
 151 aGluLeuAlaAspThrPheLysAspTyrAlaValSerThrValProValAlaAspG1 171
 524 TGAATCTTGCTGATACTTTTAAAGACTACGAGTCTTACAGTCCCTGTGGCCGATTC 465
 171 yLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuLeuAlaIleGlySerSe 191
 464 TTTGCAATTAAGAGTCTTCGACGATGGCGGACCCCACTGATTCATGCAATAGGTCACG 405
 191 rGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyrAspLy 211
 404 CGAATCTGACAGAGGCCCTCAAGATCATGCAACATGAGTACGATGATCTTATGACAA 345
 211 sLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysG1 231
 344 GCTCACTGATCCCGACGACATGGCGGCACTGCATATATCTAAATATCCCCAGCAAGG 285
 231 yHisValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLysValTyrGluLy 251
 284 GCATGCTCTGTCGACCGAACCCGAGAGAGTACCCGAGAGGCGCAAGGTCTATGAGAA 225
 251 sLeuLysAspHisMetLeuLeuProValSerMetSerGluLeuGluLysValAspGlyLe 271
 224 ACTCAGGACCATCTACTGATCCCTGTGAGCACTCGAGATGGAAAGGTGGACGGCIT 165
 271 uLeuThrCysCysSerValLeuLeuAsnLysLysValAspSer 285
 164 GCTCACTGCTGCTCGTCTTTTATTAAACAGAGATAGACTCC 122

RESULT 8
 BY742062
 DEFINITION
 BY742062
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 656)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schenbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
 Hume, D.A., Quackenbush, J., Schrim, L.M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusio, V.,
 Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustinich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
 Kawaji, H., Kawasawa, Y., Kezierski, R.M., King, B.L., Konggaya, A.,
 Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
 Maltais, L., Marchionni, L., McKenzie, B., Miki, H., Nagashima, T.,
 Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
 Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
 Sandelin, A., Schneider, C., Semp, C.A., Setou, M., Shimada, K.,
 Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
 Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
 Wells, C., Wilming, L.G., Wyszew-Boris, A., Yanagisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
 Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imorani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
 Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,

Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 22354683
 PUBMED
 12466851
 COMMENT
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp/
 URL: http://genome.gsc.riken.go.jp/
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,
 Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,
 Imorani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,
 Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,
 Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,
 Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,
 Takeda, Y., Waki, K., Watanabe, M., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 RNA was provided by Dr. Akira Nakagawara (Div. of Biochemistry,
 Chiba Cancer Center Research Institute, 666-2 Nitona, Chuo-ku,
 Chiba, 260-8717 Japan) whose assistance is gratefully acknowledged.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.

FEATURES
 Location/Qualifiers
 1..656
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="7120491N07"
 /sex="mixed"
 /tissue_type="sympathetic ganglion"
 /dev_stage="2 days neonate"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 2 days neonate
 sympathetic ganglion"
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5'
 GAGAGAGAGAGATCCAGAGCTCTTTTTTTTTTTTTTTTNN 3'] cDNA was
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. Second strand cDNA was prepared with the
 primer adapter of sequence [5'
 GAGAGAGAGATTCGAGTTAAATTAATTAATCCCCCCCCCC 3']. cDNA
 was cleaved with XhoI and BamHI. Vector: a modified
 pBluescript KS(+) after bulk excision from Lambda F1C I.
 Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

-RNA was provided by Akira Nakagawara, Div. of Biochemistry, Chiba Cancer Research Institute, 666-2 Nitona, Chuoh-ku, Chiba, 260-8717 Japan, whose assistance we gratefully acknowledge."

ORIGIN

Alignment Scores:		9,45e-114	Length:	656
Pred. No.:	62	1056.00	Matches:	209
Score:	3	98.62%	Conservative:	6
Percent Similarity:	82	95.87%	Mismatches:	3
Best Local Similarity:	63	72.63%	Indels:	1
Query Match:	13		Gaps:	0
DB:				
US-09-889-733B-2 (1-285) x BY742062 (1-656)				
QY	62	ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValala	81	
DB	3	GTGGTGCAGCTGCCCGCCGACGAGACCTGCCGACTGCGTGTTCGTGAGACGTGCC	62	
QY	82	ValValCysGluGluThrAlaLeuLeuThrArgProGluValAlaProSerArgArgLysGlu	101	
DB	63	GTGCTGTGGAGAGACGCGCCTCATCCCGCCGCGGCGCCCGCCGAGGAGAG	122	
QY	102	ValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnLeuValGluMetLysAsp	121	
DB	123	GTTGACATGATCAAGAGCGCTTTGGAAACCTTCAGCTCAACATAGTAGAGATGAAGAT	182	
QY	122	GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheValGly	141	
DB	183	GAATATGCAACTTTGGATGTGGGACGTCCTATTCACAGGACGAGATTTTGTGGC	242	
QY	142	LeuSerLysArgThrAsnGlnArgGlyAlaGluLeuAlaAspThrPheLysAspTyr	161	
DB	243	CTTTCCAAAGAACAAATCAACAGAGTGTGAAATCTTGGCTGATACCTTTAAGACTAC	302	
QY	162	AlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAla	181	
DB	303	GCAGTCTCTACAGTCCCTGTGGCGGATCTTTGCAATTAAGAGATTTCTGCAGCATGCC	362	
QY	182	GlyProAsnLeuLeuAlaLeuGlySerSerGluSerAlaGlnLysAlaLeuLysMet	201	
DB	363	GGACCCAACTGATTCAGTAGGTCCAGGATCTGCACAGAGCCCTCAAGATCATG	422	
QY	202	GlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspLeuAlaAsn	221	
DB	423	CAACAGATGAGTACCATCGTTATGACAAAGCTCACTGTACCCGACACATGCCGCCAAC	482	
QY	222	CysIleTyrLeuAsnLeuProAsnLysGlyHisValLeuLeuHisArgThrProGluGlu	241	
DB	483	TGCATATATCTAATATCCCAACCAAGGCGATGCTTGTGTCACCG-ACCCAGAGAG	541	
QY	242	TyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuLeuProValSer	261	
DB	542	TACCCAGAAAGCGCAAGGTCTATGAGAACTCAAGGACCATCTACTGATCCCTGTGAC	601	
QY	262	MetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuLeu	279	
DB	602	AATCGGAGATGGAAGAGGTGGACGCTGCTCCCTGCTCCGTTTNTTA	655	

RESULT 9
CF548604
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CF548604
AGENCOURT 15597764 NICHG XGC Brnl Xenopus laevis cDNA clone
INAME:7018702 5', mRNA sequence.
CF548604
CF548604.1 GI:34885436
EST.
Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus.

REFERENCE

1 (bases 1 to 860)
NIH-MGC <http://mgi.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaps-i@mail.nih.gov
Tissue Procurement:
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LIA414738 row: 0 column: 20
High quality sequence stop: 715.
Location/Qualifiers
1. 860
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:7018702"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHG XGC Brnl"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5 Kb. Constructed by Life Technologies. Note: This is a Xenopus Gene Collection (XGC) library."

FEATURES

source
1. 860
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:7018702"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHG XGC Brnl"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5 Kb. Constructed by Life Technologies. Note: This is a Xenopus Gene Collection (XGC) library."

ORIGIN

Alignment Scores:		4,7e-104	Length:	860
Pred. No.:	77	975.50	Matches:	201
Score:	77	86.87%	Conservative:	24
Percent Similarity:	77	77.61%	Mismatches:	32
Best Local Similarity:	77	67.09%	Indels:	4
Query Match:	14		Gaps:	1
DB:				
US-09-889-733B-2 (1-285) x CF548604 (1-860)				
QY	1	MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla	20	
DB	77	ATGGCTGACATGCTCTGAGGCTTTGGAGATGACCCACCTTAGTTCGTGCA	133	
QY	21	LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluValAspVal	40	
DB	134	ATTCTGACTCTTTTATGTTCAAGAGCTCTGAGGATAGAGAAAGCAGTGTGTGATATC	193	
QY	41	AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu	60	
DB	194	TCCCTGGCTAGGAACACACATGACGATGATGTTGAGGTGCTGAAAGATAAATACTGGGGCTG	253	
QY	61	GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal	80	
DB	254	AGTGTGATAGACTTCCAGCAGATGAATCCCTCCAGATGGCCCTTTGTGGAAGATGTG	313	
QY	81	AlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLys	100	
DB	314	GCTGTGTGTGTGAGAACTCCCTGATTACACAGACCTGGGGCAGCTAGCAGAGGAAA	373	
QY	101	GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnLeuValGluMetLys	120	
DB	374	GAGACTGAAGCTGTGAGAAAGCACTGCAGATCTTCAGCTAATGTGTTGAATGACA	433	
QY	121	AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal	140	
DB	434	GATGAAATGCTACATTAGATGAGGAGATGTTTGTGTTTACGGAAGAAATTTCTTTGTG	493	
QY	141	GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuAlaAspThrPheLysAsp	160	

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BY754930
BY754930.1 GI:27187153
EST.
Mus musculus (house mouse)
Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1. (bases 1 to 698)

REFERENCE
AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Oatono, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gotohori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, I.,
Gariboldi, M., Giesi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustingich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McGenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okado, T., Pavan, W.J., Perte, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reid, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilting, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Waki, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Ito, H., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.go.jp
URL: http://genome.gsc.riken.go.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,
Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,
Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,
Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,
Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,
Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,
Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtractions of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

TITLE

JOURNAL
MEDLINE
PUBMED
COMMENT

Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
FEATURES
source
Location/Qualifiers
1. .698
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="2860038F15"
/cell_type="8 cells"
/dev_stage="8 cells embryo"
/clone_lib="RIKEN full-length enriched, 8 cells embryo"
ORIGIN
Alignment Scores: Length: 698
Pred. No.: 1.32e-100 Matches: 183
Score: 945.00 Conservative: 5
Percent Similarity: 97.41% Mismatches: 5
Best Local Similarity: 94.82% Indels: 0
Query Match: 64.99% Gaps: 0
DB: 13
US-09-889-733B-2 (1-285) x BY754930 (1-698)
QY 93 ProGlyAlaProSerArgArgLysGluValAspMetMetLysGluAlaLeuGluLysLeu 112
Db 2 CCCGGGGCGCCAGCCGAGAGAGAGGTTGACATGATGAAGAGGCTTTGGAAAAAATT 61
QY 113 GlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeuAspGlyGlyAspValLeu 132
Db 62 CAGTCTCAACATAGTAGAGATGAAGATGAAGATCAACTTTGGATGTGGACGCTCTTA 121
QY 133 PheThrGlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnAArgGlyAlaGlu 152
Db 122 TTCACGGCAGAGAAATTTTTTGTGGGCTTTCCAAAAAGAACAAATCAACGAGGTGCTGAA 181
QY 153 IleLeuAlaAspThrPheLysAspTyrAlaValSerThrValProValAlaAspGlyLeu 172
Db 182 ATCTGTGGTGTACTTTTAAAGGACTACGAGTCTCTACAGTCCCTGTGGCGGATTTCTTG 241
QY 173 HisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySerSerGlu 192
Db 242 CATTTAAGAGTTTCTGCAGCATGCGCGGACCAACCTGATTGCAATAGGGTCCAGCGAA 301
QY 193 SerAlaGlnLysAlaLeuLysIleMetGlnMetSerAspHisArgTyrAspLysLeu 212
Db 302 TCTGCACAGAGGGCCCTCAAGATCATCAACAGATGAGTGACCATCGTTTATGACAAGCTC 361
QY 213 ThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHis 232
Db 362 ACTGTACCCGAGCAGCATGGCCGCCCACTGCATATATTCTAATATATCCCGACCAAGGGCAT 421
QY 233 ValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLysValTyrGluLysLeu 252
Db 422 GTCTTGTCTGCACCGAACCCCAAGAGAGTACCCAGAAAGCGCAAAAGGTCTATGAGAAATC 481
QY 253 LysAspHisMetLeuIleProValSerMetSerGluLeuGluLysValAspGlyLeuLeu 272
Db 482 AAGGACCATCTACTGATCCCTGTGAGCAACTCGGAGATGGAAAGGTGGACGGCTTGCTC 541
QY 273 ThrCysCysSerValLeuIleAsnLysLysValAspSer 285
Db 542 ACCTGCTGCTCGTTTTTTATTAAACAGAGATAGACTCC 580
RESULT 12
CD751679
LOCUS
DEFINITION
AGENCY: 14624576 NCI CGAP ZEMB2 Danio rerio cDNA clone
IMAGE: 696f621 5', mRNA sequence.
ACCESSION
CD751679
VERSION
CD751679.1 GI:32335966
KEYWORDS
EST.
Danio rerio (zebrafish)

ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 776)
NHL-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics / NIH
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: gcapbs-x@mail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14604 row: n column: 12
High quality sequence stop: 698.

FEATURES
source
1..776
Location/Qualifiers
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:6964621"
/tissue_type="embryo"
/lab_host="DH10B (Ti-resistant)"
/clone_lib="NCI CGAP Z5mb2"
/notes="Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dT. Average insert
size 2 kb. Constructed by J. Wang (Research Genetics,
Invitrogen Corp) from tissue donated by L. Zon (Harvard
University). Note: this is a NCI CGAP Library."

ORIGIN
Alignment Scores:
Pred. No.: 3,61e-98 Length: 776
Score: 925.00 Matches: 181
Percent Similarity: 83.20% Conservative: 27
Best Local Similarity: 72.40% Mismatches: 42
Query Match: 63.62% Indels: 0
DB: 14 Gaps: 0

US-09-889-733B-2 (1-285) x CD751679 (1-776)

Qy 8 SerAlaPheGlyArgAlaThrHisAlaValValArgAlaLeuProGluSerLeuCysGln 27
Db 26 TCCGACITCGCAAGTACACGCACGCGATCATGCGCGATGCGCGGTCTCTGAGCAC 85
Qy 28 HisAlaLeuArgSerAlaLysGlyGluValValAspValAlaArgAlaGluArgGlnHis 47
Db 86 GCCGCGCTGCGCTCCGATAATGTGTCCTCCCGTGGATCCGCTCGGGGTGCAGCGGAGTTT 145
Qy 48 GlnLeuTyrValGlyValLeuGlySerLysLeuGlnValValGluLeuProAla 67
Db 146 GACACATATGTCTGCTGCTGAGGACCGGTCTCGGCTGAGGTGGTGGTGGTGGTGGT 205
Qy 68 AspGluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGluThr 87
Db 206 GACGAGGAGCTGCCGACTGCGGTGTCGTGAGGACACGCGCGTGGTGGTGGGAGCAGC 265
Qy 88 AlaLeuLeuThrArgProGlyValAlaProSerArgArgLysGluValAspMetLysGlu 107
Db 266 GCGCTCATCACAGACCCGAGCCGCGAGCGCGAGGTGGCGAGCGGTGGCGATGAGGAT 325
Qy 108 AlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAenAlaThrLeuAsp 127
Db 326 GCGCTGACTGAGCTCGGCTCGAATTTGGAGATGAATGATGAATCAGCCAGCATGGAT 385
Qy 128 GlyGlyAspValLeuPheThrGlyArgGluPheValGlyLeuSerLysArgThrAsn 147

Db 386 CGAGGAGACGCTCTGTTCACAGGAAAGAGAGTTCTTTGTGGGATATCAAGAGACCAAT 445
Qy 148 GlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrValPro 167
Db 446 CAGAGAGAGCAGAGATCTGCGCCACACCTTAAAGGATTAAGCTGTGTCCACTATCCCT 505
Qy 168 ValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAla 187
Db 506 GTAGAGGAGGCTCTCCATCTCAAGAGGTTCTGCAAGTATGGCTGCTCAAAACCTCATGCC 565
Qy 188 IleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHis 207
Db 566 ATCGGCTCCAGTGATGCGCCACAGAGCCCTGAAGGTAATACAGCAGATGATGATTGT 625
Qy 208 ArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIle 227
Db 626 AAATATGAGAAGCTGACGTTCTCTGATGATCGGCTGCAAACTGTGTATCATCAAGCTG 685
Qy 228 ProtenLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLys 247
Db 686 CCGGGAAAGAGAGCGTTCTCTGCTGCTGACACCTGAAGTGTTCCTCGAGAGCGCCNAG 745
Qy 248 ValTyrGluLysLeuLysAspHisMetLeu 257
Db 746 GTGTTCGAGAGCGTGAAGAGTCACATGCTC 775

RESULT 13
BB611533
LOCUS
DEFINITION BB611533 RIKEN full-length enriched, 13 days embryo head Mus
ACCESSION BB611533
VERSION BB611533.1 GI:16452732
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 731)
AUTHORS Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,
Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,
Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,
Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/
Itoh, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Carninci, P., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1637-1630 (2000)
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, F.,
Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA

```

1. .897
source
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7037029"
/tissue_type="whole body"
/lab_host="DH10B"
/clone_lib="NIH ZGC 7"
/note="Vector: pExpress; Site 1: NotI; Site 2: EcoRV;
Bulk tissue was collected from a whole adult individual
from the Tuebingen strain. 1st strand cDNA was primed with
a Not I - oligo(dT) primer, double-stranded cDNA was
cloned into the Not I and EcoRV sites of pExpress-1.
Library was size-selected for >1 kb fragments and
normalized. A non-normalized version of this library is
also available (NIH ZGC 10). Library was constructed by
Open Biosystems (Huntsville, AL)"

```

ORIGIN

Alignment Scores:	
Pred. No.:	3.15e-94
Score:	892.50
Percent Similarity:	90.51%
Best Local Similarity:	68.01%
Query Match:	61.38%
DB:	14
	Length:
	Matches:
	Conservative:
	Mismatches:
	Indels:
	Gaps:

Db	623	GCTGATACCTTTAAAGGACTACGAGCTCTCTACAGTCCCTGTGGCGGATCTTTCCATTTA	682
Qy	175	LysSerPheCysSerMet-AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAl	194
Db	683	AAGACTTTCTGCAGCATGCCCGGACCCCAACTGATTGCAATAGGTCCAGCGAATCTTG	742
Qy	194	aGln---LysAlaLeuLysIleMetGlnGln-MetSerAspHisArgTyrAspLysLeuT	213
Db	743	CACAGGAAGGCGCTCAAGAATCATTTGCACAGATTGAGTGACCATGGTATGACAAAGCTCA	802
Qy	213	hrVal-ProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHis	232
Db	803	CTGTTACCGGAGACATTGCCGCGCACTGCATAATATTAAATATCCAGCAAGGCA---	859
Qy	233	ValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLysValTyrGluLysLeu	252
Db	860	TGCTTTCTGGACCGAAC-CCAGAGAGTAGTACCCAAAAGCGCAAGGCGC-TATGAGAACTTA	917

Search completed: June 8, 2004, 01:05:11
 Job time : 2536 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 7, 2004, 23:13:51 ; Search time 89 Seconds
(without alignments)
1777.090 Million cell updates/sec

Title: US-09-889-733B-2
Perfect score: 1454
Sequence: 1 MAGLHPSAFGRATHAVRA.....EKVDLLTCCSLVINKKVD5 285

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/usptp.spool.p/US09889733/runat_07062004_083529_8000/app_query.fasta_1.455
-Da=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09889733 @CGN_1.1.105 @runat_07062004_083529_8000 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	645.5	44.4	1397	4	US-09-023-655-37
2	301	20.7	2598	4	Sequence 2768, Ap
3	301	20.7	2790	4	Sequence 2864, Ap
C 4	233	16.0	239	4	Sequence 404, App
C 5	233	16.0	239	4	Sequence 404, App
C 6	233	16.0	239	4	Sequence 404, App
C 7	233	16.0	239	4	Sequence 404, App
C 8	233	16.0	239	4	Sequence 404, App
C 9	198	13.6	526	4	Sequence 12670, A
10	146.5	10.1	513	4	Sequence 370, App
C 11	122.5	8.4	4403765	3	US-09-103-840A-2
C 12	122.5	8.4	4411529	3	US-09-103-840A-1

13	113.5	7.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
14	113.5	7.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
15	102	7.0	1390	3	US-08-605-150A-7	Sequence 7, Appli
16	101.5	7.0	1354	3	US-08-605-150A-1	Sequence 1, Appli
17	97.5	6.7	1263	4	US-09-134-001C-1590	Sequence 1590, Ap
18	97.5	6.7	1440	4	US-09-252-991A-4726	Sequence 4726, Ap
19	97.5	6.7	2780	4	US-09-620-312B-358	Sequence 358, App
20	97.5	6.7	3060	4	US-09-252-991A-4686	Sequence 4686, Ap
21	97.5	6.7	3456	4	US-09-252-991A-4757	Sequence 4757, Ap
C 22	93	6.4	1708	4	US-09-919-172-30	Sequence 30, Appli
C 23	93	6.4	1709	3	US-09-010-398-2	Sequence 2, Appli
C 24	93	6.4	1709	3	US-09-366-260-2	Sequence 2, Appli
C 25	93	6.4	1709	3	US-09-153-804-11	Sequence 11, Appli
26	92.5	6.4	1490	3	US-08-605-150A-5	Sequence 5, Appli
27	90.5	6.2	1464	3	US-08-605-150A-3	Sequence 3, Appli
28	90.5	6.2	2807	4	US-09-453-702B-51	Sequence 51, Appli
29	90	6.2	1068	4	US-09-328-352-434	Sequence 434, App
30	89.5	6.2	1545	2	US-08-628-039-1	Sequence 1, Appli
31	89.5	6.2	1545	2	US-08-628-039-5	Sequence 5, Appli
32	89.5	6.2	1545	2	US-08-628-039-8	Sequence 7, Appli
33	89.5	6.2	1545	2	US-08-628-039-8	Sequence 8, Appli
34	89.5	6.2	1545	2	US-08-673-388-1	Sequence 1, Appli
35	89.5	6.2	1545	2	US-08-673-388-7	Sequence 7, Appli
36	89.5	6.2	1545	2	US-08-673-388-8	Sequence 8, Appli
37	89.5	6.2	1545	2	US-08-673-388-8	Sequence 1, Appli
38	89.5	6.2	1545	2	US-08-614-877-1	Sequence 1, Appli
39	89.5	6.2	1545	2	US-08-614-877-5	Sequence 5, Appli
40	89.5	6.2	1545	2	US-08-614-877-7	Sequence 7, Appli
41	89.5	6.2	1545	2	US-08-614-877-8	Sequence 8, Appli
42	89.5	6.2	1545	3	US-08-912-205-1	Sequence 1, Appli
43	89.5	6.2	1545	3	US-08-912-205-5	Sequence 5, Appli
44	89.5	6.2	1545	3	US-08-912-205-7	Sequence 7, Appli
45	89.5	6.2	1545	3	US-08-912-205-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-023-655-37
; Sequence 37, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023.655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1397 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: HMCINOT01
CLONE: 003437
US-09-023-655-37

Alignment Scores:
Pred. No.: 1,68e-72
Score: 645.50
Percent Similarity: 68.75%
Best Local Similarity: 51.84%
Query Match: 44.39%
DB: 4

US-09-889-733B-2 (1-285) x US-09-023-655-37 (1-1397)

Qy 3 GlyLeuGlyHisPro---SerAlaPheGlyArgAlaThrHisAlaValValArgAlaLeu 21
Db 332 GGGATGGGAGCGCGGGAGGGCTGGCGCTGCTCCATGCCCTGATCCGGGGAGTC 391
Qy 22 ProGluSerLeuGlyHisAlaLeuArgSerAlaGlyGlyGluGluValAspValAla 41
Db 392 CCAGAGAGCTGGCGTGGGGAGAGTGGGGGCTGGCTCCCGCTGCTGATCTGGCC 451
Qy 42 ArgAlaGluArgGlnHisGlnLeuTyValGlyValLeuGlySerLeuGlyLeuGln 61
Db 452 AAAGCTCAAGGGAGCAGCGGGTCTGGAGGTAACCTGAGGCAACGACTGGGGCTACAG 511
Qy 62 ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAla 81
Db 512 CTGCTAGAACCTGCCTAGAGAGTCTATCCGCTGGAGCCGCTGCTGGCGACAGGCC 571
Qy 82 ValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLysGlu 101
Db 572 GTGATCAAGGGAGCAGCGGCCCTAATCAGCGGCGCTGGAGCCCGCTGCTAGGCCAGAG 631
Qy 102 ValAspMetLysGluAlaLeuGluLysLeuGlnLeuAsnLeuValGluMetLysAsp 121
Db 632 GTGATGAGTCCCAAGCCCTCAAGACCTGGGGCTCCGATTGTGGAAATAGAGAC 691
Qy 122 GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGly 141
Db 692 GAGAACGCGAGCTGGATGGCACTGACGTTCTTTCACCGCGCGGAGTTTTCGTAGGC 751
Qy 142 LeuSerLysArgThrAsnGlnArgGlyAla-GluLeuLeuAlaAspThrPheLysAspTy 161
Db 752 CTCTCCAAATGGACCAATCACCGAGGCGCTGGAGATCGTGGCGGACACGTTCCGGGACTT 811
Qy 161 rAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAl 181
Db 812 CGCGCTCTCCACTGTGCAGTCTGGGTCTCCCTCCACTGCGCGGTCTCTGCGGCATGGG 871
Qy 181 aGlyProAsnLeuLeuAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMe 201
Db 872 GGGACCTCGCACTGTGTGGCAGGAGCAGCGAGCGCTGCCCAAAAGGCTGTCCGGGCAAT 931
Qy 201 tGlnGlnMetSerAspHisArgTyAspLysLeuThr-ValProAspAspIleAlaAlaA 221
Db 932 GCGAGTGTGACAGATACCCATATGCTCTCCCTGACCCCTCCAGATGACGCGCTGCTG 991
Qy 221 snCysIleTyLeu-----AsnIleProAsnLysGlyHisValLeuLeuHisArgThrP 239
Db 992 ACTGTCTCTTCTGCTGGTGTGCTGTGGTGGCCCTTCTCTCTCCGACCGTGGAG 1051
Qy 239 roGluGlu-----TyProGluSerAlaLysValTyGluLysLeuLysAspHisMetL 257

Db 1052 GTGGGATCTGTGNCNCCCAAAACAGCAGGAGGCACTGCAGAGCTCTTGTGATGCACCC 1111
Qy 257 euileProValSerMetSerGluLeuGluLys 267
Db 1112 TGGTACCTGTGTCTGTCTCAGACTGGAGAG 1143
RESULT 2
US-09-252-991A-2768
; Sequence 2768, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2768
; LENGTH: 2598
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2768

Alignment Scores:
Pred. No.: 5.56e-28
Score: 301.00
Percent Similarity: 48.50%
Best Local Similarity: 30.83%
Query Match: 20.70%
DB: 4

US-09-889-733B-2 (1-285) x US-09-252-991A-2768 (1-2598)

Qy 15 HisAlaValValArgAlaLeuProGluSerLeuGlyHisAlaLeuArgSerAlaLys 34
Db 320 CACATCATCTGTCGACCGCCCGCGAGCTGTGTCAGCGCTGACCTCCAGCCACCTC 379
Qy 35 GlyGluGluValAspValAlaAargAlaGluArgGlnHisGlnLeuTyValGlyValLeu 54
Db 380 GGCAAGCCG---GACTACGCCAAGCCCTGGAGCAGCAACAGCCCTACATCCGCGCTG 436
Qy 55 GlySerLysLeuGlyLeuGlnValValGluLeuProAlaAspGluSerLeuProAspCys 74
Db 437 ---CAGACCTGCGAGCTGGACATCACCTGCTGCGCCCGCGAGAACGCTTCCCGACTCG 493
Qy 75 ValPheValGluAspValAlaValValCysGluGluThrAlaLeuLeuThrArgProGly 94
Db 494 GTGTTCTGTCGAGGACCCGCTGCTGTCACCTGCGCTGCGCATCATCACCGCCCGCGC 553
Qy 95 AlaProSerArgArgLysGluValAspMetLysGluAlaLeuGluLysLeuGlnLeu 114
Db 554 GCCGAATCGCGCGCGGCGAGACCGAGATCATCGAGGAAACCGTCGACGCTTCTATCCG 613
Qy 115 AsnIleValGluMetLysAspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThr 134
Db 614 GCGAAGTTCGAGCATCGAGCACCCTGCGCGCTGGAGCGCGGACATCATCATGTGTC 673
Qy 135 GlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuLeu 154
Db 674 GCGCAGCACTTCTACATCGCGAATCGCGCCGACCAACCGCGAGGCGCGCGGAGATG 733
Qy 155 AlaAspThrPheLysAspTyAlaVal-----SerThrValProValAlaAspGlyLeu 172
Db 734 ATCGCGATCTCGAGAAACATGCGCTCAGCGCTGCTGTCGCTGCGCTGAAAGATCTG 793
Qy 173 HisLeuLysSer---PheCysSerMetAlaGlyProAsnLeuLeuAlaIleGlySerSer 191
Db 794 CACCTGAACCGGGTCTGCTTACCTGGAACACAAACCTGCTGCGCGCGCGGAGTTC 853

QY 192 GluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyrAspLys 211
Db 854 GTCAGC-----AAGCGGAGTTCACGACTTCAACATCATCGAG 892
QY 212 LeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGly 231
Db 893 ATCCCGAGAGAGAGTCTACGCGCACTCATCTGGGTCAAC-----GAA 940
QY 232 HisValLeuLeuHisArgThrProGlnGluTyrProGluSerAlaLysValTyrGluLys 251
Db 941 AGGGTGATCATG-----CCCGCGGCTATCCCGGACCCCGGAGAAATCGCCCGC 991
QY 252 LeuLysAspHisMetLeuIleProValSerMetSerGluLeuGluLysValAspGlyLeu 271
Db 992 CTC---GGCTACCGGGTGTATCGAGTGGACACCTCCGAATATCGAAGATCGACGGCGC 1048
QY 272 LeuThrCysCysSerVal 277
Db 1049 GTCAGTTGCATGTCGCTG 1066
RESULT 3
US-09-252-991A-2864
; Sequence 2864, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107195.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2864
; LENGTH: 2790
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2864
Alignment Scores:
Pred. No.: 6,24e-28 Length: 2790
Score: 301.00 Matches: 82
Percent Similarity: 48.50% Conservative: 47
Best Local Similarity: 30.83% Mismatches: 117
Query Match: 20.70% Indels: 20
DB: 4 Gaps: 8
US-09-889-733B-2 (1-285) x US-09-252-991A-2864 (1-2790)
QY 15 HisAlaValAlaArgAlaLeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLys 34
Db 2035 CACATCATCGCTCGACGCGCGCGCGCGCGCTGGTGTGACGGCTGACCTCCAGCCACCTC 2094
QY 35 GlyGluGluValAspValAlaAlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeu 54
Db 2095 GGCAAGCCG---GACTACCGCCAGGCGCTGGAGCAGCAACGCTTACATCGCGCGCTTG 2151
QY 55 GlySerLysLeuGlyLeuGlnValValGluLeuProAlaAspGluSerLeuProAspCys 74
Db 2152 ---CAGACCTCGGACGTGGACATCACCTGCTGCGCGCGGAGCGGTTCCTCCGACTCG 2208
QY 75 ValPheValGluAspValAlaValAlaValCysGluGluThrAlaLeuIleThrArgProGly 94
Db 2209 GTGTTCTGTCGAGGACCGCGGTGCTTCCTCACCTCGCGCTGCGGCTATCATCACCGCGCGC 2268
QY 95 AlaProSerArgArgLysGluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeu 114
Db 2269 GCCCAATCGCGCGCGCGGAGACCGGAGATCATCGAGAAACCGTGCAGCGCTTCTATCCG 2328

QY 115 AsnIleValGluMetLysAspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThr 134
Db 2329 GGCAAGTTCGCGCATCGAGGACCGCGGACCGGTGGAGCGCGGACATCATGATGTC 2388
QY 135 GlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeu 154
Db 2389 GGGACCACTTCTACATCGCGGAAATCGCGCCGACCAACCGCCGCGCGCGCGCATG 2448
QY 155 AlaAspThrPheLysAspTyrAlaVal-----SerThrValProValAlaAspGlyLeu 172
Db 2449 ATCGGATCTCGGAGAAACATGCGCTCAGCGGCTCGGTGTGCGCTGGAAAGGTCCTG 2508
QY 173 HisLeuLysSer---PheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySerSer 191
Db 2509 CACCTGAAGACCGGGTCTCGCTACCTGGAAACACACCAACCTGTGCGCGCGCGCGAGTTC 2568
QY 192 GluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyrAspLys 211
Db 2569 GTCAGC-----AAGCGGAGTTCAGGACTTCAACATCATCGAG 2607
QY 212 LeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGly 231
Db 2608 ATCCCGAAGAGGAGTCTACGCGCGCACTGATCTGGGTCAAC-----GAA 2655
QY 232 HisValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLysValTyrGluLys 251
Db 2656 AGGTGATCATG-----CCCGCGGCTATCCCGGACCCCGGAGAAAGATCGCCCGC 2706
QY 252 LeuLysAspHisMetLeuIleProValSerMetSerGluLeuGluLysValAspGlyLeu 271
Db 2707 CTC---GGCTACCGGGTGTATCGAGTGGACACCTCCGAATATCGAAGATCGACGGCGC 2763
QY 272 LeuThrCysCysSerVal 277
Db 2764 GTCAGTTGCATGTCGCTG 2781
RESULT 4
US-09-702-705-404/C
; Sequence 404, Application US/09702705
; Patent No. 8504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-404
Alignment Scores:
Pred. No.: 5,85e-21 Length: 239
Score: 233.00 Matches: 49
Percent Similarity: 78.48% Conservative: 13
Best Local Similarity: 62.03% Mismatches: 17
Query Match: 16.02% Indels: 0
DB: 4 Gaps: 0
US-09-889-733B-2 (1-285) x US-09-702-705-404 (1-239)

```
Qy 62 ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAla 81
Db 239 CTGCTAGAACTGCCACCTGAGGAGTCAATGCCCGCTGGGACCGCTGCTGGCGACACGGCC 180
Qy 82 ValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGlu 101
Db 179 GTGATCAAGGGGACACGGCCCTAATCAGCGGCCCTGGAGCCCGCTCGTAGGCCAGAG 120
Qy 102 ValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAsp 121
Db 119 GTCGATGGAGTCCGCAAGCCCTGCAAGACCTGGGGCTCCGAATTTGGAATAGGAGAC 60
Qy 122 GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal 140
Db 59 GAGAACGGGACGCTGGATGGCACTGACGTTCTTTCACCGCCGGGAGTTTTCGTA 3

RESULT 5
US-09-736-457-404/c
; Sequence 404, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736.457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-404

Alignment Scores:
Pred. No.: 5,85e-21 Length: 239
Score: 233.00 Matches: 49
Percent Similarity: 78.48% Conservative: 13
Best Local Similarity: 62.03% Mismatches: 17
Query Match: 16.02% Indels: 0
DB: 4 Gaps: 0

US-09-889-733B-2 (1-285) x US-09-736-457-404 (1-239)
Qy 62 ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAla 81
Db 239 CTGCTAGAACTGCCACCTGAGGAGTCAATGCCCGCTGGGACCGCTGCTGGCGACACGGCC 180
Qy 82 ValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGlu 101
Db 179 GTGATCAAGGGGACACGGCCCTAATCAGCGGCCCTGGAGCCCGCTCGTAGGCCAGAG 120
Qy 102 ValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAsp 121
Db 119 GTCGATGGAGTCCGCAAGCCCTGCAAGACCTGGGGCTCCGAATTTGGAATAGGAGAC 60
Qy 122 GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal 140
Db 59 GAGAACGGGACGCTGGATGGCACTGACGTTCTTTCACCGCCGGGAGTTTTCGTA 3

RESULT 6
US-09-614-124B-404/c
; Sequence 404, Application US/09614124B
```

```
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614.124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-404

Alignment Scores:
Pred. No.: 5,85e-21 Length: 239
Score: 233.00 Matches: 49
Percent Similarity: 78.48% Conservative: 13
Best Local Similarity: 62.03% Mismatches: 17
Query Match: 16.02% Indels: 0
DB: 4 Gaps: 0

US-09-889-733B-2 (1-285) x US-09-614-124B-404 (1-239)
Qy 62 ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAla 81
Db 239 CTGCTAGAACTGCCACCTGAGGAGTCAATGCCCGCTGGGACCGCTGCTGGCGACACGGCC 180
Qy 82 ValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGlu 101
Db 179 GTGATCAAGGGGACACGGCCCTAATCAGCGGCCCTGGAGCCCGCTCGTAGGCCAGAG 120
Qy 102 ValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAsp 121
Db 119 GTCGATGGAGTCCGCAAGCCCTGCAAGACCTGGGGCTCCGAATTTGGAATAGGAGAC 60
Qy 122 GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal 140
Db 59 GAGAACGGGACGCTGGATGGCACTGACGTTCTTTCACCGCCGGGAGTTTTCGTA 3

RESULT 7
US-09-671-325-404/c
; Sequence 404, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671.325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
```



```
/ NAME/KEY: CDS
/ LOCATION: 277..513
/ NAME/KEY: sig_peptide
/ LOCATION: 277..399
/ OTHER INFORMATION: Von Heijne matrix
/ OTHER INFORMATION: score 3.59899990463257
/ OTHER INFORMATION: seq AGAGLPALDLAKA/QR
US-09-621-976-370

Alignment Scores:
Pred. No.: 2,36e-09 Length: 513
Score: 146.50 Matches: 31
Percent Similarity: 64.79% Conservative: 15
Best Local Similarity: 43.66% Mismatches: 24
Query Match: 10.08% Indels: 1
DB: 4 Gaps: 1

US-09-889-733B-2 (1-285) x US-09-621-976-370 (1-513)
QY 3 GlyLeuGlyHisPro---SerAlaPheGlyArgAlaThrHisAlaValAlaValAlaLeu 21
DB 274 GGGATGGGAGCCCGGGGAGGGGCTGGCGCGCTGCTCCCATGCCCTGATCCGGGGAGTC 333
QY 22 ProGluSerLeuGlyGlnHisAlaLeuArgSerAlaLysGlyGluGluValAlaValAla 41
DB 334 CCAGAGAGCCTGGCTCGGGGAGAGGTGGCGGGGCTGGCCCTCCCGCTCTGGATCTGGCC 393
QY 42 ArgAlaGluArgGlnHisGlnLeuTyValGlyValLeuGlySerLysLeuGlyLeuGln 61
DB 394 AAAGTCAAGGAGCAGCGGGGTGCTGGAGGTAACTGAGGCAACGACTGGGCTACAG 453
QY 62 ValValGluLeuProAlaAspGluSerLeuPro 72
DB 454 CTGCTAGACTGCCACCTGAGGAGTCATTCGCG 486

RESULT 11
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 6.57 Length: 4403765
Score: 122.50 Matches: 63
Percent Similarity: 36.53% Conservative: 36
Best Local Similarity: 23.25% Mismatches: 107
Query Match: 8.43% Indels: 65
DB: 3 Gaps: 10

US-09-889-733B-2 (1-285) x US-09-103-840A-2 (1-4403765)
QY 37 GluValAspValAlaArgAla-----GluArgGlnHisGlnLeuTyValGlyValLeu 54
DB 374 GGGATGGGAGCCCGGGGAGGGGCTGGCGCGCTGCTCCCATGCCCTGATCCGGGGAGTC 333
```

```
Db 2591873 GACGTCCAAAGTCGCGCAGCAGCAGTGGGAGCAGCCTCCACGACCTATCTT----- 2591823
QY 55 GlySerLysLeuGlyLeuGlnValValGluLeuProAlaAspGluSerLeuProAspCys 74
Db 2591822 -----CGGTAGGCCACAGCGTGGATCTGATCGAGCCCATTTCCGGGTATACCGACATG 2591769
QY 75 ValPheValGluAspValAlaValValCysGluGluThrAlaLeuLeuThrArgProGly 94
Db 2591768 GTGTACACGCCCAACGGTGGTTCATCGCGCAGCAGATCGCGTGGTCCCGGTTCCGG 2591709
QY 95 AlaProSerArgLysGluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeu 114
Db 2591708 TTCCCGCAGACGAGCTGGTGAAGTCTAGAGCCTATGACAGCTGGATCTCTCGGTCCGATAT 2591649
QY 115 AsnIleValGluMetLysAspGluAsnAlaThrLeuAspGly---GlyAspValLeuPhe 133
Db 2591648 CGCCCGGTGACCCCGCCGACGCTAAC-----GAGGACAGCGGCGCTGCTGATG 2591598
QY 134 ThrGlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGly---AlaGlu 152
Db 2591597 GTTGGCGAAAGGGTGTGGCGGGCTACGGCTTTCGACAGACAGCAGCGCGCACACCGCGAA 2591538
QY 153 IleLeuAla----- 155
Db 2591537 ATCGCGCGCGTGGTCTGCGGGTGGTCTCCCTCGAGTGGTGGTCCACCGGTTCTAT 2591478
QY 156 -----AspThr-----PheLysAspTyArgAlaValSerThrValProValAla 169
Db 2591477 CACCTGACACCGCGTGGCGGTGCTCGACGACACACAGCTGCTACTACCGCGCGCG 2591418
QY 170 AspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGly 189
Db 2591417 TTCAGTACGGCAGCGCAGGAAACAGTGTGCGGGCTGTTCCCGACCGCATTTGGTTCGCG 2591358
QY 190 SerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTy 209
Db 2591357 AGTGGCGCGCGTTCGTGTCGAGCTCAACGCGTCTCT----- 2591319
QY 210 AspLysLeuThrValProAspAspIleAlaAlaAsnCysIleTyLeuAsnIleProAsn 229
Db 2591318 GACGGTCTGAACGTAGTGTTCGCGTGGCGGCATGGGTTTTCGCGCGCAGGTACCGCGCA 2591259
QY 230 LysGlyHisValLeuLeuHisArgThrProGluGluTyProGluSerAlaLysValTy 249
Db 2591258 GCGGCG-----TTC 2591250
QY 250 GluLysLeuLysAspHisMetLeuLeuProValSerMetSerGluLeuLysValAsp 269
Db 2591249 GAGCGG-----GTCGGTGTGATCTGTCCGAGCTGTCTCAAGGCGCGC 2591208
QY 270 GlyLeuLeuThrCysCysSerValLeuLeuAsn 280
Db 2591207 GGTTCGCTCAAGTCTGACCGCTGGAGATACAC 2591175

RESULT 12
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
```



```
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 658 Length: 4411529
Score: 12250 Matches: 63
Percent Similarity: 36.53% Conservative: 36
Best Local Similarity: 23.25% Mismatches: 107
Query Match: 8.43% Indels: 65
DB: 3 Gaps: 10

US-09-889-733B-2 (1-285) x US-09-103-840A-1 (1-4411529)
QY 37 GluValAspValAlaArgAla-----GluArgGlnHisGlnLeuTy-ValGlyValLeu 54
Db 2596063 GAGCTCCAAAGTCGCGAAGCAGATGGGAGACCTCCACACACATATCTT----- 2596013
QY 55 GlySerLeuGlyLeuGlnValValGluLeuProAlaAspGluSerLeuProAspCys 74
Db 2596012 -----CGCTAGGCCACAGCGTGGATCTGATCGAGCCCATTTCCGGGTACCGACATG 2595959
QY 75 ValPheValGluAspValAlaValValCysGluGluThrAlaLeuLeuThrArgProGly 94
Db 2595958 GTGTACACGCCCAACAGCTGGTTCATCGCGCAGCATCGCGTGTGCGCCGGTTCGGG 2595899
QY 95 AlaProSerArgArgLysGluValAlaAspMetMetLysGluAlaLeuGlnLysLeuGlnLeu 114
Db 2595898 TTCCCGAAGAGCTGGTGGATCTAGAGCCTATGCCAGCTGGATGCTCGCTCGATAT 2595839
QY 115 AsnIleValGluMetLysAspGluAsnAlaThrLeuAspGly---GlyAspValLeuPhe 133
Db 2595838 CGCCCGGTGACACCCGCCACGCTCAAC-----GAGGACAGCGGCACCTGCTGATG 2595788
QY 134 ThrGlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGly---AlaGlu 152
Db 2595787 GTTGGCGAAGAGGTGTGGCGGCTACGCTTTCGCACAGACAGCGCGCACACGCCGAA 2595728
QY 153 IleLeuAla----- 155
Db 2595727 ATCGCGCGGTGCTTGGTCTGCGCGGTGCTCCCTCGAGTGGTGCACCCAGGTTCTAT 2595668
QY 156 -----AspThr-----PheLysAspTyAlaValSerThrValProValAla 169
Db 2595667 CACCTGGACACCCGCGTGGCGGTGCTCGACGACACACAGCTACCTACCGCGCGCGG 2595608
QY 170 AspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGly 189
Db 2595607 TTCAGTACGCGCAGCGAGAACAGTTGTCGCGCTGTTCCCGCAGCCGATTTGGTGTGCGC 2595548
QY 190 SerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyr 209
Db 2595547 AGTCCGACGCGTGGTGTTCGAGCTCAACGCGTCTCT----- 2595509
QY 210 AspLysLeuThrValProAspAspIleAlaAlaAsnCysIleTyLeuAsnIleProAsn 229
Db 2595508 GACGGTCTGAACGTAGTGTTCGCGTGGCGCCATGGGTTTTCGCGCGAGTTACCGCA 2595449
QY 230 LysGlyHisValLeuLeuHisArgThrProGluGluTy-ProGluSerAlaLysValTyr 249
Db 2595448 GCCGCG-----TTC 2595440
QY 250 GluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGlnLysValAsp 269
Db 2595439 GAGCGG-----GTCGGTGTGATCTGTCCGAGTGTCTCAAGGCGCGC 2595398
QY 270 GlyLeuLeuThrCysCysSerValLeuIleAsn 280
Db 2595397 GGTTCGCTCAAGTCTGCAAGCTGAGATACAC 2595365

RESULT 13
US-09-103-840A-2
; Sequence 2, Application US/09103840A
```

```
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
```

```
Alignment Scores:
Pred. No.: 93.1 Length: 4403765
Score: 113.50 Matches: 65
Percent Similarity: 34.83% Conservative: 36
Best Local Similarity: 22.41% Mismatches: 98
Query Match: 7.81% Indels: 91
DB: 3 Gaps: 11

US-09-889-733B-2 (1-285) x US-09-103-840A-2 (1-4403765)
QY 59 GlyLeuGlnValValGluLeuProAlaAsp----- 68
Db 1117574 GCGATGACCTTCAACGAGCTCCGTCGACACGCGGACCGGTGCTGGTGGTTCGCT 1117633
QY 69 -----GluSerLeuProAspCysValPheValGluAsp 79
Db 1117634 ATGCACCATGCGGAGACTTCGTCATTGACCGGTGCGCAACTGGTGTTCACCCGCGAC 1117693
QY 80 ValAlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArg 99
Db 1117694 TCGTCGATGATGATCGGCGCGGTGCTGATCCGTCGCTGGCATTCACGGCAGCGGTG 1117753
QY 100 LysGluValAspMet-----MetLys 106
Db 1117754 CGCGAAGCGTGGCTGACCGACCTCATCTATGCTCATCACCCGCGGTTCACCGGTGCGG 1117813
QY 107 GluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeu 126
Db 1117814 CGTGCTATGAA-----TCGCGCAGCGCTCCGTC 1117843
QY 127 AspGlyGlyAspValLeuPheThrGlyArgGluPhePhe---ValGlyLeuSerLysArg 145
Db 1117844 GAGGTGGCGAGCTGTTGCTCGCCCGGTGCTGCTGCTGCGAGTGGCGGAGCGG 1117903
QY 146 ThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAsp-----Tyr 161
Db 1117904 ACTACACGACGAGCGCGGAGCATTTGCGCGACCTTTTGACGATGATCTTCGCGCAT 1117963
QY 162 AlaValSerThrValProValAlaAsp-----GlyLeuHisLeuLysSerPheCys 178
Db 1117964 ACCGTGCTCGCGTGGCGGATCGCTCAGCAGCGCGGCAAAATGATCTCGACACGCGTGTG 1118023
QY 179 SerMetAlaGly----- 182
Db 1118024 ACCATGGTCGACACCGATACGATGATGTGATGATGATGATGATGATGATGATGATGAT 1118083
QY 183 -----ProAsnLeuIleAlaIleGlySer-----SerGlu 192
Db 1118084 TTCACGATCCAGCGCACACCGCAGCGGTGACCATTCGCGGATGCGGCCCGCTTCGCGGAG 1118143
```

```
QY 193 SerAlaGlnLysAlaLeuLysLeuMetSerAspHisArgTyrAspLysLeu 212
Db 1118144 CGGGTCCCAAGCGATGGGAATC-----GACAGCTG 1118176

QY 213 ThrVal-----ProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIlePro 228
Db 1118177 CGGGTAATTCATACCGGAATGACCCCGTCTCGCTGAAACGGAACAGTGGGACGAGCGC 1118236

QY 229 AsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLysVal 248
Db 1118237 AACAAACAGTTGGCGTTGGCCCGCGGTGCTGTTGTCGCTACGAGCGCAACGTCACAGCC 1118296

QY 249 TyrGluLysLeuLysAspHis-----MetLeuIleProValSerMetSerGluLeuGlu 266
Db 1118297 AACGCCCGCTGCAGGACGGCGCATCGAAGTGTCTTACCATCGCGCGTCCGAATTGGGT 1118356

QY 267 LysValaspGlyLeuLeuThrCysCysSer 276
Db 1118357 ACCGGCCGTGGCGGCGCCCGCTGCATGTCC 1118386

RESULT 14
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 93.4 Length: 4411529
Score: 113.50 Matches: 65
Percent Similarity: 34.83% Conservative: 36
Best Local Similarity: 22.41% Mismatches: 98
Query Match: 7.81% Indels: 91
DB: 3 Gaps: 11

US-09-889-733B-2 (1-285) x US-09-103-840A-1 (1-4411529)
QY 59 GlyLeuGlnValValGluLeuProAlaAsp----- 68
Db 1117552 GGCATGACCTTCACGAGCTCCGTCGACACGCGACCGACGTGTCGTGTTGGT 1117611

QY 69 -----GluSerLeuProAspCysValPheValGluAsp 79
Db 1117612 ATGCACCATGGCGGAGACTTCGTCTATTAGCCGTTGCCGACCTGCTGTTCCCGGAC 1117671

QY 80 ValAlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArg 99
Db 1117672 TCGTCGATATGATCGGGCGCGGGTGGTGTGTCGCTCGCTGGCATTACGGGACGGGTG 1117731

QY 100 LysGluValAspMet-----MetLys 106
Db 1117732 CCGGAAGCGTCCGTCACCGACCTCATCTATGCTCATCCCGCGGTTCCCGGTGTGGCG 1117791

QY 107 GluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeu 126
Db 1117792 CGTGCTATGAA-----TCGCGCACCGCTCCGGTC 1117821
```

ATTORNEY/AGENT INFORMATION:
NAME: Szekeres, Gabor L.
REGISTRATION NUMBER: 28,675
REFERENCE/DOCKET NUMBER: 542-04-PA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-854-5502
TELEFAX: 714-854-4897
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1390 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: Cuphea lanceolata
IMMEDIATE SOURCE: ZAP cDNA library
LIBRARY: C1GPDH30
CLONE: C1GPDH30
FEATURE:
NAME/KEY: CDS
LOCATION: 34..1149

US-08-605-150A-7

Alignment Scores:
Pred. No.: 0.00587 Length: 1390
Score: 102.00 Matches: 70
Percent Similarity: 37.50% Conservative: 38
Best Local Similarity: 24.31% Mismatches: 100
Query Match: 7.02% Indels: 80
DB: 3 Gaps: 16

US-09-889-733B-2 (1-285) x US-08-605-150A-7 (1-1390)

QY 10 PheGlyArgAlaThrHisAlaValValArgAlaLeuProGluSerLeu---CysGlnHis 28
DB 4 TTCGGCAGAGTTTCTTCACCCCTCGCAATGGCTCCCTCTGAGCTCAACTGCACCCAC 63
QY 29 AlaLeuArgSerAlaGlyGluGluValAaspValAlaArgAlaGluArgGlnHisGln 48
DB 64 CAGAACCCACATTCAGCGGTTACGACGGACCCAGATCGAGG----- 105
QY 49 LeuTyrValGlyValLeuGlySer---LysLeuGlyLeuGlnValValGluLeuProAla 67
DB 106 -----GTCCCGTTGTCGGTAGTGAACCTCGGCGAGTGTGCGTCCCAAGCTATTGCT 159
QY 68 AspGluSerLeuProAspCysValPheValGluAaspValAlaVal---ValCysGluGlu 86
DB 160 TCCATACCTCAAGCTTCCATCTTTTCATGATGAAGTGAGAATCTGGGTATTTGAGGAG 219
QY 87 ThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLysGluValAaspMetMetLys 106
DB 220 ACT-----CTACCGAGCGGAGAGCTTACTGATGTCATCAAC 258
QY 107 GluAlaLeuGlu-----LysLeuGlnLeuAsnIleValGlu 118
DB 259 CAGACCAATGAATAATTTAAGTATCTCCCGGAATTAAGCTCGGTAGGAATGTTGTTGCA 318
QY 119 MetLysAasp---GluAsnAlaThrLeuAaspGlyGlyAaspValLeuPhe---ThrGlyArg 136
DB 319 GATCCAGACCTCGAAAACCGAGTTAAGATGCAAAATATGCTCGTGTGTCACACCGCAT 378
QY 137 GluPhePheValGlyLeuSerLysArg-----ThrAsnGlnArgGlyAlaGluIle 153
DB 379 CAGTTTCATGAGGGCATCTGCAAAAGACTCGTAGGGAATAATACAGAGGACACAGGCT 438
QY 154 LeuAlaAaspThrPheLysAaspTyrAlaValSerThrValProValAlaAaspGlyLeuHis 173
DB 439 CTTCC-----CTTATAAAGGGCATGGAG 462
QY 174 LeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySerSerGluSer 193

DB 463 GTCAAA-----ATGAGGGGCTTGCATGATCTCG----- 492
QY 194 AlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyrAspLysLeuThr 213
DB 493 -----AGCCTA 498
QY 214 ValProAspAspIleAlaAlaIleCysIleTyrLeu-----AsnIleProAsnLys 230
DB 499 ATCTCTGATCTTCTCGGATCAACTGCTGTCTCTAATCGGGGCAACATCGCTAATGAG 558
QY 231 GlyHisValLeuLeuHisArgThrProGluGlu-----TyrProGluSerAlaLys 247
DB 559 -----ATTGCTGTTGAGAAATTCAGTGAAGCGACAGTCGGGTTTCAGAGAAATACAGAT 612
QY 248 ValTyrGluLysLeuLysAspHisMetLeuIleProValSerMet---SerGluLeuGlu 266
DB 613 ATTGCGAGAAATGGGTTTCAGCTCTTTAGCACTCCGTTACTTTCATGCTCTCAGCTGTGAA 672
QY 267 LysValAaspGlyLeuLeuThrCys 274
DB 673 GATGTTGAGGAGTAGAACTTTGT 696

Search completed: June 8, 2004, 03:30:29
Job time : 8710 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 7, 2004, 23:21:06 ; Search time 389 Seconds

(without alignments)

3342.336 Million cell updates/sec

Title: US-09-889-733B-2

Perfect score: 1454

Sequence: 1 MAGLGHPSAFGRATHAVRA.....EKVDGLLTCCSVLKNKVDV 285

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delpop 6.0 , Delext 7.0	

Searched: 2995936 seqs, 2280999010 residues

Total number of hits satisfying chosen parameters: 5991872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xdp
-Q=/cgn2_1/USPTO.spool_p/US09889733/runat_07062004_083529_8024/app_query_faasta_1.455
-DB=Published Applications NA -QWMI=fastcap -SUFFIX=p2n.rnnp -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09889733 @cgn1_1_723 @runat_07062004_083529_8024
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100
-LONGLOG -DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1:	/cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:
2:	/cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:
3:	/cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:
4:	/cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:
5:	/cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:
6:	/cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:
7:	/cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:
8:	/cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:
9:	/cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:
10:	/cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:
11:	/cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:
12:	/cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:
13:	/cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:
14:	/cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:
15:	/cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:
16:	/cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:
17:	/cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:
18:	/cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:
19:	/cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

RESULT 1

US-10-450-826-9
; Sequence 9, Application US/10450826
; Publication No. US20040101818A1
; GENERAL INFORMATION:
; APPLICANT: Ji, Darren
; APPLICANT: Axelrod, Douglas W.
; APPLICANT: Cook, Jonathan S.
; APPLICANT: Jaiswal, Neelam
; APPLICANT: Eistein, Richard
; APPLICANT: Houghton, Adam
; APPLICANT: Mertiz, Lawrence
; TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiation
; FILE REFERENCE: 044921-5039-WO/US/10/450,826
; CURRENT APPLICATION NUMBER: US/10/450,826
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 60/255,882
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/285,691
; PRIOR FILING DATE: 2001-04-24
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 9

ALIGNMENTS

1	1439	99.0	1633	17	US-10-450-826-9	Sequence 9, Appli
2	1439	99.0	4208	17	US-10-468-123-22	Sequence 22, Appli
3	1379	94.8	3008	15	US-10-205-219-83	Sequence 83, Appli
4	1200	82.5	3778	15	US-10-252-157-328	Sequence 328, App
5	1064	73.2	3651	9	US-09-925-300-512	Sequence 512, App
6	645.5	44.4	1397	17	US-10-641-643-37	Sequence 37, Appli
7	551.5	37.9	770	9	US-09-910-943-9	Sequence 9, Appli
8	386	26.5	472	10	US-09-918-995-9117	Sequence 9117, Ap
9	382	26.3	452	10	US-09-918-995-14871	Sequence 14871, A
10	342	23.5	403	10	US-09-918-995-19254	Sequence 19254, A
11	282.5	19.4	9025608	15	US-10-156-761-1	Sequence 1, Appli
12	281	19.3	774	15	US-10-156-761-1679	Sequence 1679, Ap
13	266	18.3	332	9	US-09-783-590-5309	Sequence 5309, Ap
14	265.5	18.3	83945	17	US-10-450-826-8	Sequence 8, Appli
15	233	16.0	239	9	US-09-736-457-404	Sequence 404, App
16	233	16.0	239	9	US-09-902-341-404	Sequence 404, App
17	233	16.0	239	9	US-09-849-626-404	Sequence 404, App
18	233	16.0	239	10	US-09-476-300-404	Sequence 404, App
19	233	16.0	239	13	US-10-283-017-404	Sequence 404, App
20	233	16.0	239	15	US-10-017-754-404	Sequence 404, App
21	233	16.0	239	15	US-10-113-872-404	Sequence 404, App
22	204	14.0	3673778	15	US-10-312-841-1	Sequence 1, Appli
23	198	13.6	3189	10	US-09-814-353-20553	Sequence 20553, A
24	165	11.3	3673778	15	US-10-312-841-2	Sequence 2, Appli
25	153.5	10.6	475	10	US-09-918-995-16593	Sequence 16593, A
26	117	8.0	1377	13	US-10-282-122A-17396	Sequence 17396, A
27	107.5	7.4	1218	15	US-10-238-075-714	Sequence 714, App
28	107.5	7.4	1349	16	US-10-369-493-38269	Sequence 38269, A
29	107.5	7.4	1785	16	US-10-369-493-38603	Sequence 38603, A
30	107.5	7.4	1821	16	US-10-369-493-35077	Sequence 35077, A
31	107.5	7.4	3943	16	US-10-398-221-3727	Sequence 3727, Ap
32	107.5	7.4	6164	15	US-10-238-075-705	Sequence 705, App
33	107.5	7.4	6425	15	US-10-085-959-217	Sequence 217, App
34	105.5	7.3	1059	16	US-10-369-493-37878	Sequence 37878, A
35	103.5	7.1	1227	15	US-10-156-761-2307	Sequence 2307, Ap
36	103.5	7.1	9025608	15	US-10-156-761-1	Sequence 1, Appli
37	102.5	7.0	1827	10	US-09-738-269-40	Sequence 40, Appli
38	102.5	7.0	1827	14	US-10-023-437-40	Sequence 40, Appli
39	102.5	7.0	684707	16	US-10-398-221-9	Sequence 9, Appli
40	102.5	7.0	3011208	16	US-10-398-221-2058	Sequence 2058, Ap
41	102	7.0	163	15	US-10-029-386-26599	Sequence 26599, A
42	102	7.0	528	15	US-10-029-386-12999	Sequence 12999, A
43	102	7.0	1230	16	US-10-369-493-40383	Sequence 40383, A
44	101.5	7.0	135638	15	US-10-314-657-1	Sequence 1, Appli
45	101	6.9	2832	13	US-10-282-122A-31868	Sequence 31868, A

US-09-889-733B-2 (1-285) x US-10-468-125-22 (1-4208)

```

Qy 261 SerMetSerGluLeuGluValAspGlyLeuLeuThrCysCysSerValLeuLeuAsn 280
Db 1103 AGCATGCTCTGAAGTGAAGAAGTGGATGGGCTGCTCACCTGCTGCTCAGTTTAAATAC 1162
Qy 281 LysLysValAspSer 285
Db 1163 AACAAAGTAGACTCC 1177

RESULT 2
US-10-468-125-22
; Sequence 22, Application US/10468125
; Publication No. US20040082061A1
; GENERAL INFORMATION:
; APPLICANT: ASTROWOFF, Anna
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: DING, Li
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: FORSYTHE, Ian J.
; APPLICANT: GRIFFIN, Jennifer J.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: LU, Yan
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: RING, Huijun Z.
; APPLICANT: SANJANWALA, Madhusudan
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: WARREN, Bridget A.
; APPLICANT: XU, Yuming
; APPLICANT: YUE, Henry
; APPLICANT: ZEBARJADIAN, Yeganeh
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0363 USN
; CURRENT APPLICATION NUMBER: US/10/468,125
; CURRENT FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: PCT/US02/04918
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,643
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/271,332
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/276,767
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: US 60/282,077
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/285,447
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/287,060
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/288,543
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 22
; LENGTH: 4208
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7671089CB1
US-10-468-125-22

Alignment Scores:
Pred. No.: 9,59e-177 Length: 4208
Score: 1439.00 Matches: 283
Percent Similarity: 99.65% Conservative: 1
Best Local Similarity: 99.30% Mismatches: 1
Query Match: 98.97% Indels: 0
DB: 17 Gaps: 0

```

QY 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
Db 368 ATGGCCGGGCTGGCCACCCCGCCCTTCGCGCGGCCACCCACCGCGTGTGGCGGCG 427
QY 21 LeuProGluSerLeuGlyHisAlaLeuArgSerAlaGlyGluGluValVal 40
Db 428 CTACCCGAGTGGCTGGCGGAGCAGCCGCTGAGAGCGCCAAAGCGGCGAGGAGTGGACGTC 487
QY 41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
Db 488 GCCCGCGGGAACGCGACAGCAGCAGCTCTACGTGGCGGTGGCGGAGCAAGCTGGGCGTG 547
QY 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
Db 548 CAGGTGGTGGAGTCCCGCGGCGGAGAGCCCTCCGAGTGGCTCTTCGTGGAGAGCGTG 607
QY 81 AlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLys 100
Db 608 GCCGTGGTGTGGAGGAGAGCGCCCTCATCCCGAGCCCGGGCGCGGAGGAGAG 667
QY 101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnLysValGluMetLys 120
Db 668 GAGGTGTGACATGATGAAGAAGCATTAGAAAATTCAGCTCAATATATAGTAGAGATGAAA 727
QY 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal 140
Db 728 GATGAATATGCAACTTTAGATGGCGGAGATGTTTATTCACAGGCGAGAGATTTTGTG 787
QY 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuLeuAlaAspThrPheLysAsp 160
Db 788 GGCCCTTCCAAAAGGACAAATCAACGAGGTGCTGAATCTTGGCTGATATCTTTAAGGAC 847
QY 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
Db 848 TATGCAGTCTCCACAGTGCAGTGCAGATGGGTTCGCAATTTGAAGAGTTCTTCGACGATG 907
QY 181 AlaGlyProAsnLeuLeuAlaLeuGlySerGluSerAlaGlnLysAlaLeuLysLys 200
Db 908 GCTGGGCTTAACCTGATCGCAATTTGGTCTAGTGAATCTGCACAGAGCGCCCTTAAGATC 967
QY 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspLysAla 220
Db 968 ATGCAACAGATGAGTGACCCCGTACGCAAACTCACTGTGCTGTGATGACATAGACGA 1027
QY 221 AsnCysLysLeuAsnLysProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
Db 1028 AACTGTATATATCTAAATATCCCAACAAAGGCGACGCTTGTCTCACCGAACCCCGGAA 1087
QY 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuLeuProVal 260
Db 1088 GAGTATCCAGAAAGTGCAAGGTTTATGAGAAACTGAAGGACCATATGCTATCCCCGTG 1147
QY 261 SerMetSerGluLeuGluLysValAspGlyLeuThrCysCysSerValLeuLeuLeu 280
Db 1148 AGCATGTCTGAATCGAAAGGTGATGGCTGCTCACCTGCTGCTCAAGTTTAAATTAAC 1207
QY 281 LysLysValAspSer 285
Db 1208 AAGAAAGTAGACTCC 1222

RESULT 3
US-10-205-219-83
; Sequence 83, Application US/10205219
; Publication No. US20030138603A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pincock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018200
; CURRENT APPLICATION NUMBER: US/10/205,219

; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 83
; LENGTH: 3008
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: N-G, N-G dimethylarginine dimethylaminohydrolase
; NAME/KEY: misc feature
; LOCATION: 2411, 2412
; OTHER INFORMATION: n is a or g or c or t
US-10-205-219-83
Alignment Scores:
Pred. No.: 3 98e-169 Length: 3008
Score: 1379.00 Matches: 268
Percent Similarity: 96.84% Conservative: 8
Best Local Similarity: 94.04% Mismatches: 9
Query Match: 94.84% Indels: 0
DB: 15 Gaps: 0
US-09-889-733B-2 (1-285) X US-10-205-219-83 (1-3008)
QY 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
Db 432 ATGGCCGGGCTGGCCACCCCGCCCTTCGCGCGGCCACCCACCGCGTGTGGCGGCT 491
QY 21 LeuProGluSerLeuGlyHisAlaLeuArgSerAlaLysGlyGluGluValAspVal 40
Db 492 CGCCCGAGTTCCTGTGGCGGAGCGGTGAGCGGCTCCCGAGGCGGAGGTGGATTC 551
QY 41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
Db 552 GCTCGCGCTGAGCGCGGAGCAGCAGCTCTACGTGGCGGTGTGGCGGAGCAAGCTGGGCGTG 611
QY 61 GluValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
Db 612 CAGGTGGTGGAGTGGCGGCGGAGCGCTTCGCTGAGTGGCTGTTCGTGGAGGAGCGTG 671
QY 81 AlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLys 100
Db 672 GCCGTGGTGTGGAGGAGACCGCCCTCATCCCGCGCGGCGGCGCTAGCCGACGAGAG 731
QY 101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnLysValGluMetLys 120
Db 732 GAGGTGGACATGATGAAGAGCGCTTTGGAAAACTTCAGCTCAACATAGTAGAGATGAAA 791
QY 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal 140
Db 792 GATGAATATCAACCTTAGATGTGGGAGCTCTATTACAGCGCAGAGAGTTTGTG 851
QY 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuLeuAlaAspThrPheLysAsp 160
Db 852 GGCCCTTCCAAAAGGACAAATCAACGAGGTGCTGAGATCTTGGCTGATATCTTTCAAGGAC 911
QY 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
Db 912 TAGCAGTTTCCACAGTCCCGCGGCGATCTTTGCAITTTAAAGAGTTTCTCGACGATG 971
QY 181 AlaGlyProAsnLeuLeuAlaLeuGlySerSerGluSerAlaGlnLysAlaLeuLysLys 200
Db 972 GCTGGCCCCAACCTGATCGCAATAGGCTCCAGTGAATCTGCGCAGAGAGGCCCTCAAGATC 1031
QY 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspLysAla 220
Db 1032 ATGCAACAGATGATGACCCCGCTTATGCAAGCTCAGTGTAGCGGAGGAGTGGCGGCC 1091
QY 221 AsnCysLysLeuLeuAsnLysProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240

```

Db 1092 AACTGTATATATTAAATATATCCAGCAAGGGCATGCTTGTGTCACCGAACCCAGAA 1151
Qy 241 GlnTyrProGluSerAlaLeuValTyrGluLysLeuLysAspHisMetLeuLeuProVal 260
Db 1152 GAGTACCAGAACCGCAAGGTTTATGAGAGCTCAAGACCATCTACTGATCCCTGTG 1211
Qy 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuLeuAsn 280
Db 1212 AGCAATCTGAGATGGAAAGGTGAGCGGCTTGCTCACCTGCTGCTCCGTTTTATTAAAC 1271
Qy 281 LysLeuValAspSer 285
Db 1272 AAGAAGACAGACTCT 1286

RESULT 4
US-10-252-157-328
; Sequence 328, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Paris, Mary
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 328
; LENGTH: 3778
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030190640A1 198087.1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1083-1106
; OTHER INFORMATION: a, t, c, g, or other
US-10-252-157-328

Alignment Scores:
Pred. No.: 1,43e-145 Length: 3778
Score: 1200.00 Matches: 238
Percent Similarity: 99.58% Conservative: 0
Best Local Similarity: 99.58% Mismatches: 0
Query Match: 82.53% Indels: 1
DB: 15 Gaps: 0

US-09-889-733B-2 (1-285) x US-10-252-157-328 (1-3778)
Qy 48 GlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValGluLeuProAla 67
Db 2 CAGCTCTACGTGGCGTGTGGCGAGCAAGCTGGGCTGCGAGGTGGAGCTGCCGGCC 61
Qy 68 AspGluSerLeuProAspCysValPheValGluAspValAlaValCysGluThr 87
Db 62 GACGAGAGCCTTCGAGCTGCGTCTTCGTGGAGACGTGGCGGTGGTGCGAGAGACG 121
Qy 88 AlaLeuLeuThrArgProGlyAlaProSerArgArgLysGluValAspMetMetLysGlu 107
Db 122 GCCCTCATCCCGGCGCGCGGCGAGGAGGAGGTGACATATGAAGAA 181
Qy 108 AlaLeuGluLysLeuGlnLeuAsnLeuValGluMetLysAspGluAsnAlaThrLeuAsp 127
Db 182 GCATTAGAAACCTTCAGCTCAATATAGTACAGATGAAAGATGAAATGCAACTTTAGAT 241
Qy 128 GlyGlyAspValLeuPheThrGlyArgGluPheValGlyLeuSerLysArgThrAsn 147
Db 242 GCGCGAGATGTTTATTCACAGCGAGAAATTTTGTGGCCCTTCCAAAGACAAAT 301
Qy 148 -GlnArgGlyAlaGluLeuLeuAlaAspThrPhelLysAspTyrAlaValSerThrValPr 167

```

```

Db 302 CCACAGAGGTGCTGAAATCTTTGGCTGATACTTTTAAGGACTATGCAGTCTCCAGTCCC 361
Qy 167 oValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuLeuAl 187
Db 362 AGTGACAGATGGTTCATTTGAGAGGTTTCTGACGATGGCTGGGCTTACCTGATGCG 421
Qy 187 alleGlySerGluSerAlaGlnLysAlaLeuLysLeuMetGlnGlnMetSerAspHi 207
Db 422 AATTCGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATCATGCAACAGATGAGTGACCA 481
Qy 207 sArgTyrAspLysLeuThrValProAspAspLeuAlaLeuAsnLysTyrLeuAsnIl 227
Db 482 CCGCTACGACAACTCACTGTCCTGATGATAGCAGCAAACTGATATATCTAAATAT 541
Qy 227 eProAsnLysGlyHisValLeuLeuHisArgThrProGluLutyrProGluSerAlaLy 247
Db 542 CCCCACAAAGGGCACGCTTGTCTCACCGAACCCCGGAGAGTATCCAGAAAGTGCAAA 601
Qy 247 sValTyrGluLysLeuLysAspHisMetLeuLeuProValSerMetSerGluLeuGluLy 267
Db 602 GGTTCATGAGAACTGAAGGACCATATGTCATCCCGTGAGCATGTCTGAACCTGAAAA 661
Qy 267 sValAspGlyLeuLeuThrCysCysSerValLeuLeuAsnLysLysValAspSer 285
Db 662 GGTGGATGGCTGCTCACTGCTGCTCAGTTTAAATTAACAAGAAAGTAGACTCC 716

RESULT 5
US-09-925-300-512
; Sequence 512, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen.
; APPLICANT: Steve Ruben.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 512
; LENGTH: 3651
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1283)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3641)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3650)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-512

Alignment Scores:
Pred. No.: 8.2e-128 Length: 3651
Score: 1064.00 Matches: 212
Percent Similarity: 99.53% Conservative: 0
Best Local Similarity: 99.53% Mismatches: 1
Query Match: 73.18% Indels: 1
DB: 9 Gaps: 0

US-09-889-733B-2 (1-285) x US-09-925-300-512 (1-3651)
Qy 73 AspCysValPheValGluAspValAlaValCysGluGlnThrAlaLeuLeuThrArg 92
Db 4 GACTGGGTCTTCGTGGAGGACGTGGCGGTGGTGGAGGACGCGCCCTCATCCCCGA 63

```


Db 992 ACTGCTCTTTCTGCTGGTTGCTGGTGTGCCCCCTTCTCTCTGACCGTGGAG 1051
Qy 239 roGluGlu-----Ty-ProGluSerAlaIysValTyrGluIysLeuIysAspHisMetL 257
Db 1052 GTGGGGATCTGTGTCNCCCAACAGCAGGAGGACCTGCAGAGGCTCTCTGATGTACCC 1111
Qy 257 euileProValSerMetSerGluLeuGluIys 267
Db 1112 TGGTACCTGTGCTGCTCAGAACTGGAGAAG 1143

RESULT 7

US-09-910-943-9

; Sequence 9, Application US/09910943

; Patent No. US20020081610A1

; GENERAL INFORMATION:

; APPLICANT: Hemmati-Briavanlou, Ali

; APPLICANT: Altman, Curtis

; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression

; FILE REFERENCE: 7529/IG148US1

; CURRENT APPLICATION NUMBER: US/09/910,943

; CURRENT FILING DATE: 2001-07-23

; NUMBER OF SEQ ID NOS: 742

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 9

; LENGTH: 770

; TYPE: DNA

; ORGANISM: Xenopus laevis

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)-(770)

; OTHER INFORMATION: n may be a or g or c or t/u

US-09-910-943-9

Alignment Scores:

Pred. No.: 7,41e-62 Length: 770

Score: 551.50 Matches: 110

Percent Similarity: 74.11% Conservative: 36

Best Local Similarity: 55.84% Mismatches: 46

Query Match: 37.93% Indels: 5

DB: 9 Gaps: 1

US-09-889-733B-2 (1-285) x US-09-910-943-9 (1-770)

Qy 11 GlyArgAlaThrHisAlaValValArgAlaLeuProGluSerLeuCysGlnHisAlaLeu 30
Db 186 GGGCCCTACACTCAGCTGTGGTGGGGTGTCCCTGCTTGAAGAAAGAGCC--- 242
Qy 31 ArgSerAlaIysGlyGluValAspValAlaArgAlaGluArgGlnHisGlnLeuTyr 50
Db 243 -----AATGGTCAGGTAGACCTCGCAGCGGCTCAACGTGAGAAATGGAGTCTAC 290
Qy 51 ValGlyValLeuGlySerIysLeuGlyValValGlnValGluLeuProAlaAspGluSer 70
Db 291 TGTGGATCCTGAGACAGAACTTGGCTGAGGTGTGGAGTTGCCCCCAATGAGGAA 350
Qy 71 LeuProAspCysValPheValGluAspValAlaValCysGluGluThrAlaLeuIle 90
Db 351 CTGCCCCGGGGCAATTGATGGGACACAGCTGTAGTGATAGCAGATACAGCCCTCATC 410
Qy 91 ThrArgProGlyAlaProSerArgArgIysGluValAspMetMetIysGluAlaLeuGlu 110
Db 411 ACCCGTCCATGATACCTGCAGAGAGAGAGACTGAAGCCCTGCAAAAAATCTTTGAG 470
Qy 111 LysLeuGlnLeuAsnIleValGluMetIysAspGluAsnAlaThrLeuAspGlyGlyAsp 130
Db 471 GAGCTGAATTCGGAGTCTGGCAACTCAGTATGAAATGCCACTCTGGATGCAAGTGAT 530
Qy 131 ValLeuPheThrGlyArgGluPhePheValGlyLeuSerIysArgThrAsnGlnArgGly 150
Db 531 ATACITTTTACAGGTTTCAGAGATTTTGTAGGCTTGTCTAAATGACCAATCTTAGAGGT 590
Qy 151 AlaGluIleLeuAlaAspThrPheIysAspTyrAlaValSerThrValProValAlaAsp 170

Db 591 GCTGAATCGTGGCAAGACCTTACAGGATTATGCTGTTTCAACTGTTCTGTGCTGGG 650
Qy 171 GlyLeuHisLeuIysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySer 190
Db 651 GACATGCACTTTAAAGCTTTTGAGTATGGCAGACCTGACACCCCTGTCATAGGAAGC 710
Qy 191 SerGluSerAlaGlnIysAlaLeuIysIleMetGlnGlnMetSerAspHis 207
Db 711 ACTGATACGCGAGAAAGGCACTTGAAGACTATGGACCAACTGACCGATCAT 761

RESULT 8

US-09-918-995-9117

; Sequence 9117, Application US/09918995

; Publication No. US20030073623A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995

; PRIOR FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US/09/235,076

; NUMBER OF SEQ ID NOS: 38054

; SOFTWARE: PastSeq for Windows Version 3.0

; SEQ ID NO 9117

; LENGTH: 472

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)-(472)

; OTHER INFORMATION: n = A,T,C or G

US-09-918-995-9117

Alignment Scores:

Pred. No.: 1,51e-40 Length: 472

Score: 386.00 Matches: 83

Percent Similarity: 72.00% Conservative: 25

Best Local Similarity: 55.33% Mismatches: 30

Query Match: 26.55% Indels: 12

DB: 10 Gaps: 1

US-09-889-733B-2 (1-285) x US-09-918-995-9117 (1-472)

Qy 30 LeuArgSerAlaIysGlyGluGlu----- 37
Db 21 CTGAGACTCTCGAAGCCCGGAATTCGCGGGGAAGGTGCGGGGCTGGCCTTCCCGCT 80
Qy 38 ValAspValAlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerIys 57
Db 81 CTGGATCTGCGCAAGCTCAAAGGGAGCACGCGGGGGCTGGGAGGTAAACTGAGGCAACGA 140
Qy 58 LeuGlyLeuGlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheVal 77
Db 141 CTGGGCTTACAGCTGTAGACTGCCACCTGAGGAGGCAATTCGCCCTGGGACCGCTGCTT 200
Qy 78 GluAspValAlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSer 97
Db 201 GGCACACACGCGCCAGATCCAAAGGGAGCACGCGGCCCTTAATCACGCGGCCCTGGAGCCCCGT 260
Qy 98 ArgArgGlyValValAspMetMetIysGluAlaLeuGluIysLeuGlnLeuAsnIleVal 117
Db 261 CGTAGCCAGAGTCTGATGGAGTCCCAAGGCCCTGCTCAAGACTGGGGCTCCGAATTGTG 320
Qy 118 GluMetIysAspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGlu 137
Db 321 GAAATAGGAGACGAGAACCGCAGCTGGATGGCACTGACGCTTCTCTTCCCGCGCGGAG 380
Qy 138 PhePheValGlyLeuSerIysArgThrAsnGlnArgGlyValGluIleLeuAlaAspThr 157
Db 381 TTTTCTGAGGCTCTCCAAATGGACCAATACCGGAGGAGCTGAGATCGTGGCGGACACG 440

QY 158 PheLysAspTyrAlaValSerThrValPro 167
Db 441 TTCGGGACTTCGGCTGTACACTGTGCCA 470

RESULT 9

US-09-918-995-14871
; Sequence 14871, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14871
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(452)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-14871

Alignment Scores:
Pred. No.: 4,7e-40 Length: 452
Score: 382.00 Matches: 81
Percent Similarity: 74.81% Conservative: 20
Best Local Similarity: 60.00% Mismatches: 34
Query Match: 26.27% Indels: 0
DB: 10 Gaps: 0

US-09-889-733B-2 (1-285) x US-09-918-995-14871 (1-452)

QY 35 GlyGluValAlaValSerThrValProAlaValSerThrValGlyValLeu 54
Db 39 GGCAGCAAAACNNNAATNAGNAAGGAGCAGCGGGGCTGGAGGAAACTG 98
QY 55 GlySerLysLeuGlyLeuGlnValValGluLeuProAlaValSerThrValGlyValLeu 74
Db 99 TGGCAACAGCTGTGCTACAGCTGTAGAACTCCACCTGAGGAGTCATTCGCGCTGGGA 158
QY 75 ValPheValGluValAlaValAlaValCysGluGluThrAlaLeuLeuThrArgProGly 94
Db 159 CCGCTGCTTGGCGACACGGCGGTGATCCAAAGGGGACACGGCCCTAATCAGCGCGCTGG 218
QY 95 AlaProSerArgLysGluValAlaValMetMetLysGluAlaLeuGluLysLeuGlnLeu 114
Db 219 AGCCCGCTGTAGCGCAGAGGTGATGGAGTCCGCAAGCCCTGCAAGACCTGGGGCTC 278
QY 115 AsnLeuValGluMetLysAspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThr 134
Db 279 CGAATTGTGGAATAGAGAGCAGAACCGACGCTGGATGGCACTGACGTCTCTTCCACC 338
QY 135 GlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuLeu 154
Db 339 GCGCGGAGGATTTTCGTAGGCTCTCCAAATGGACCAATCACCAGAGAGCTGAGATCGTG 398
QY 155 AlaAspThrPheLysAspTyrAlaValSerThrValProValAla 169
Db 399 CGGACAGCTTCGGGACTTCGCGCTCTCCACTGTGCCAGTCTCG 443

RESULT 10

US-09-918-995-19254
; Sequence 19254, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19254
; LENGTH: 403
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-19254

Alignment Scores:
Pred. No.: 6,67e-35 Length: 403
Score: 342.00 Matches: 75
Percent Similarity: 71.32% Conservative: 22
Best Local Similarity: 55.15% Mismatches: 33
Query Match: 23.52% Indels: 6
DB: 10 Gaps: 1

US-09-889-733B-2 (1-285) x US-09-918-995-19254 (1-403)

QY 31 ArgSerAlaLysGlyGluValAlaValSerThrValGluLeuGlnHisGlnLeuTyr 50
Db 12 CGCTGTGATCTGCCAA-----GCTCAAGGGAGCAGCGGTGCTG 53
QY 51 ValGlyValLeuGlySerLysLeuGluValValGluLeuProAlaAspGluSer 70
Db 54 GGAGTAAACTGAGCAACACTGGGCTACAGCTGTAGAACTGCCACCTGAGGAGTCA 113
QY 71 LeuProAspCysValPheValGluAspValAlaValCysGluGluThrAlaLeuLeu 90
Db 114 TTGCGCTGGGACCGCTCTTGGGACACGGCCGTGATCCAGGGGACACGGACTAATC 173
QY 91 ThrArgProGlyAlaProSerArgLysGluValAlaValMetMetLysGluAlaLeuGlu 110
Db 174 AGCGCGCTTGGAGCCCGCTCGTAGGCCACAGGTGCGATGAGTCCGCAAGCCCTGCAA 233
QY 111 LysLeuGlnLeuAsnLeuValGluMetLysAspGluAsnAlaThrLeuAspGlyVal 130
Db 234 GACCTGGGCGCTCCGAATTTGTGAAATAGAGACGAGAACGCGCTGGATGGCACTGAC 293
QY 131 ValLeuPheThrGlyValGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGly 150
Db 294 GTTCGTTCACGGCCGGAGATTTTCAGAGCGCTCTCTAATGGACCAATCACCAGGA 353
QY 151 AlaGluLeuLeuAlaAspThrPheLysAspTyrAlaValSerThrVal 166
Db 354 GCTGAGATCGTGGCGGACACAGTTCGGGACTTCGACGCTTCACTGTG 401

RESULT 11

US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02

```
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Alignment Scores:
Pred. No.: 2,23e-20 Length: 9025608
Score: 282.50 Matches: 97
Percent Similarity: 47.65% Conservative: 45
Best Local Similarity: 32.55% Mismatches: 117
Query Match: 19.43% Indels: 39
DB: 15 Gaps: 11

US-09-889-733B-2 (1-285) x US-10-156-761-1 (1-9025608)

Qy 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArg-----12
Db 2068758 TTGTGCGCGGGGGTGCCTCGATGATGATGGCGGTACCCCTCCGAGCGAGACCG 2068817
Qy 13 -----AlaThrHisAlaValValArg-----AlaLeuProGluSerLeu 25
Db 2068818 CACGTGCCAGCAAGAGCCCTCATCGCGCCAGCCCGCCCTCCGCGAGAGGCTG 2068877
Qy 26 CysGlnHisAlaLeuArgSerAlaLysGlyGluGluValAlaArgAlaGluArg 45
Db 2068878 GTACGCACATCGAGCGCG-----CAGGTGCGATGCGCGCTCGCGTCGAG 2068925
Qy 46 GlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuValValGluLeu 65
Db 2068926 CAGTGGAGCGGTATCGGAGCGCTGCGCAGCAC---GCTGGAGACGTCGAGGTG 2068982
Qy 66 ProAlaAspGluSerLeuProAspCysValPheValGluAspValAlaValCysGlu 85
Db 2068983 GATCGCGCGGACACTGCGCGACTCGGTGTGTCGAGGACGCGGTGTCATACCG 2069042
Qy 86 GluThrAlaLeuLeuThrArgProGlyAlaProSerArgAlaGlyValAlaAspMet 105
Db 2069043 AATGTGCGCTGATCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2069102
Qy 106 LysGluAlaLeuGluLysLeuGlnLeuAsnLeuValGluMetLysAspGluAsnAlaThr 125
Db 2069103 GAGGAGCGCGTGGCGCGCTGGCTGCTCG---GTGAACCTGATCTGGAAACCGCGCAC 2069159
Qy 126 LeuAspGlyCysValLeuPheThrGlyArgGluPhePheValGlyLeuSerLysArg 145
Db 2069160 CTCGACGCGCGGACGTCCTGAAGATCGCGGACACGATCTACGTAGGCGCGCGCG 2069219
Qy 146 ThrAsnGlnArgGlyAlaGluLeuAlaAspThrPheLysAspTyrAla-----Val 163
Db 2069220 ACCNACGCGCGGGGTCCAGCAGCTGCGCGGGTCTTCGAGCGCTCGGAGCCAGGTC 2069279
Qy 164 SerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyPro 183
Db 2069280 GTCGCGCTGCGCGTACGAAAGTGTGACCTCGAAAGTCTCGTCAAC---GCGCTGCGG 2069336
Qy 184 AsnLeuLeuAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGln 203
Db 2069337 GACGACCGGTATCGGTACATCCGCTCGTGGACAGCGCGCGCTGTCTCACTCGCTTC 2069396
Qy 204 MetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsnCysIle 223
Db 2069397 CTCTCG-----GTGCGGAGGAGTCCGCTCCCATGTGTC 2069432
Qy 224 TyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrPro 243
Db 2069433 CTGCTCGC-----GGATCGAAGCTGCTG-----ATGCGGACGACGCGCCCG 2069474
```

```
Qy 244 GluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuLeuProValSerMetSer 263
Db 2069475 AAGACGGGAGCTCTCTCGCGGACCTC---GGCCACGAGCGCGTCTGTTAACATCAGC 2069531

Qy 264 GluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuLeuAsnLys 281
Db 2069532 GAGTACGAGAACTCGAAGGCTGTGTGACATGTCTCTCGGTCGCGCTGAGGAG 2069585

RESULT 12
US-10-156-761-1679
; Sequence 1679, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1679
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; NAME/KEY: CDS
; LOCATION: (1)..(774)
US-10-156-761-1679

Alignment Scores:
Pred. No.: 1,73e-26 Length: 774
Score: 281.00 Matches: 91
Percent Similarity: 49.43% Conservative: 39
Best Local Similarity: 34.60% Mismatches: 109
Query Match: 19.33% Indels: 24
DB: 15 Gaps: 9

US-09-889-733B-2 (1-285) x US-10-156-761-1679 (1-774)

Qy 21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluGluValAspVal 40
Db 43 CTCGCGAAGGGCTGTGTACGACATCGAGCGGCG-----CAGGTGATGTC 90
Qy 41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
Db 91 GGCTTCGCGTTCAGCAGTGGAGGCGGTATCGGAGCGCTGCGGACGAC---GGCTGG 147
Qy 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
Db 148 GAGACCTCGAGGTGATCGGCGGACGACTGCGCGACTCGGTGTCTGTCGAGGACGCG 207
Qy 81 AlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgAlaLys 100
Db 208 GTGTCTATGACCGCAATGTCGCCCTGTATCACCCTCCGCGCGCGCGCGCGCGCGCG 267
Qy 101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
Db 268 GAGACCGCGCGCTCGAGGAGCGCGTGGCGCGCTGGCTGCTCG---GTGAACCTGATC 324
Qy 121 AspGlnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheVal 140
Db 325 TGGGAACCGCGCACCTTCGACGCGCGCGCGCTCCTGAAGATCGCGGACACGATCTACGTA 384
Qy 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuAlaAspThrPheLysAsp 160
```



```
; TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiation
; FILE REFERENCE: 044921-5039-WO
; CURRENT APPLICATION NUMBER: US/10/450,826
; PRIORITY FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 60/255,882
; PRIORITY FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/285,691
; PRIORITY FILING DATE: 2001-04-24
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 8
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. AL078459
US-10-450-826-8

Alignment Scores:
Pred. No.: 2,648-21 Length: 83946
Score: 265.50 Matches: 77
Percent Similarity: 45.18% Conservative: 12
Best Local Similarity: 39.09% Mismatches: 53
Query Match: 18.26% Indels: 55
DB: 17 Gaps: 6

US-09-889-733B-2 (1-285) x US-10-450-826-8 (1-83946)
QY 85 GluGluThrAlaLeuLeuThrArgProGly-----94
DB 18988 GAGGAGGAGCCCTGCTTACTACGCCAAGATATCAAAACCCCTGGCGCTTACATGT 18929
QY 95 -----AlaProSerArgArgLysGluValAspMet 104
DB 18928 GGCCATTTTGATCTTACTGATAGTCTTACGACAAAGACTAGAAAATAAATAGCGGT 18869
QY 105 MetLysGluAlaLeuGluLysLeuGlnLeuAsnLleValGluMetLysAspGluAsnAla 124
DB 18868 TTTAAACCTTTAAAGAAAGAGAGACAAACACATACTTTTAAAGAGACAAACAAACCT 18809
QY 125 ThrLeuAspGlyGlyAsp-----ValLeuPhe-----Thr 134
DB 18808 TGCAATCTTCCAAAGACTCATCTTAATCAAGTTTCTTTTCTGTTCTAAGTGACT 18749
QY 135 GlyArg-----GluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyAla 151
DB 18748 GGAGGTAATAGATGGGAATTT-----GAACAGAGAAACAGCCTAAAGTGT 18704
QY 152 GluLeuLeuAlaAspThrPheLysAspTyrAlaValSerThrValProValAlaAspGly 171
DB 18703 CAGTTCCAAAGCTTGATCTCTACCACTGCGCTTCTGTTGAAAGTGTGGGAGTTA 18644
QY 172 LeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuLleAlaIleGlySerSer 191
DB 18643 CTCCTTTAGGCTGCTTTGTTTC-----18620
QY 192 GluSerAlaGlnLysAlaLeuLysIleMetGlnMetSerAspHisArgTyrAspLys 211
DB 18619 -----TCTTTATAGATCATGCAACAGATGAGTACCAACCGCTACGACAAA 18575
QY 212 LeuThrValProAspAlaAlaAsnCysIleTyrLeuAsnLleProAsnLysGly 231
DB 18574 CTCAGTGTGCTGATGATAGACAGACCAACTGTATATATCTAATATATCCCAACAAAGG 18515
QY 232 HisValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLysVal 248
DB 18514 CACGTCTTGTGTCACCGAACCCCGAAGAGTATCCAGAAAGTGCAGAAAGGTA 18464
```

RESULT 15

US-09-736-457-404/c

; Sequence 404, Application US/09736457

; Patent No. US20020168637A1

; GENERAL INFORMATION:

```
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-404

Alignment Scores:
Pred. No.: 5,288-21 Length: 239
Score: 233.00 Matches: 49
Percent Similarity: 78.48% Conservative: 13
Best Local Similarity: 62.03% Mismatches: 17
Query Match: 16.02% Indels: 0
DB: 9 Gaps: 0

US-09-889-733B-2 (1-285) x US-09-736-457-404 (1-239)
QY 62 ValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAla 81
DB 239 CTGCTAGAACTGCCACCTTGAGAGTCAATTCGCTGGGACCGCTCTTGGCGACACGCC 180
QY 82 ValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLysGlu 101
DB 179 GTGATCCAAAGGGGACACGCGCTTAATCACGGCGCTTGAGCCCGCTCTGTAGGCCAGAG 120
QY 102 ValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnLleValGluMetLysAsp 121
DB 119 GTCGATGGAGTCCGCAAGCCCTGCAGACCTGGGGCTCCGAATTTGGAATAGGAGAC 60
QY 122 GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal 140
DB 59 GAGAACGGGACGCTGGATGGCACTGACGTTCTTTCACCGCGGGAGTTTTCGTA 3

Search completed: June 8, 2004, 02:56:42
Job time : 6580 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 8, 2004, 00:22:57 ; Search time 3209 Seconds
(without alignments)
3849.412 Million cell updates/sec

Title: US-09-889-733B-2

Perfect score: 285
Sequence: 1 MAGLHPSAFGRATHAVRA.....EKVDGLTCCSVLKNKVD 285

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6934735

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2.1/USPTO spool.p/US09889733/runat.07062004.083557.8353/app_query.fasta.1.455
-DB=GenEmbl -OPWT=fastap -SUFFIX=olip2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODES=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09889733 @CNC 1.1.5600 @runat.07062004.083557.8353 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6
-Fgapext=7 -Ygapop=60 -Ygapext=60 -DELOP=6 -DELEXT=7

Database : GenEmbl:
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_ov.*
5: gb_ov.*
6: gb_ov.*
7: gb_pl.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ro.*
12: gb_ro.*
13: gb_ro.*
14: gb_ro.*
15: gb_ro.*
16: gb_ro.*
17: gb_ro.*
18: gb_ro.*
19: gb_ro.*
20: gb_ro.*
21: gb_ro.*
22: gb_ro.*
23: gb_ro.*
24: gb_ro.*
25: gb_ro.*
26: gb_ro.*
27: gb_ro.*
28: gb_ro.*

29: em_vi.*
30: em_hgt_hum.*
31: em_hgt_inv.*
32: em_hgt_inv.*
33: em_hgt_inv.*
34: em_hgt_inv.*
35: em_hgt_inv.*
36: em_hgt_inv.*
37: em_hgt_inv.*
38: em_hgt_inv.*
39: em_hgt_inv.*
40: em_hgt_inv.*
41: em_hgt_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	285	100.0	858	6	BD251986 Screen me
2	285	100.0	858	6	AX032835 Sequence
3	259	90.9	1624	9	BC033680 Homo sapi
4	259	90.9	1633	6	BD094073 Shear str
5	259	90.9	1633	9	AB001915 Homo sapi
6	259	90.9	4208	6	AX548046 Sequence
7	246	86.3	3737	9	BC043235 Homo sapi
8	184	64.6	3908	9	EX648145 Homo sapi
9	106	37.2	1120	10	BC034505 Mus muscu
10	106	37.2	3008	6	AX683109 Sequence
11	106	37.2	3008	10	D86041 Rattus norv
12	84	29.5	2895	9	AK098405 Homo sapi
13	76	26.7	104071	2	AL162260 Homo sapi
14	76	26.7	158405	2	AC019261 Homo sapi
15	76	26.7	168953	9	AL360219 Human DNA
16	69	24.2	2289	10	AK128909 Mus muscu
17	49	17.2	83946	9	HSAL23M24 Human DNA
18	48	16.8	1344	5	BC056074 Xenopus l
19	46	16.1	3420	10	BC060266 Mus muscu
20	38	13.3	186434	2	AC126412 Mus muscu
21	38	13.3	228940	2	AC126890 Rattus no
22	34	11.9	186434	2	AC126412 Mus muscu
23	34	11.9	200007	2	AC123684 Mus muscu
24	34	11.9	212111	2	AC094482 Rattus no
25	31	10.9	1150	5	AB095027 Gallus ga
26	27	9.5	420	6	AX898477 Sequence
27	27	9.5	420	6	BD034010 Sequence
28	26	9.1	110000	2	Continuation (3 of
29	26	9.1	304208	2	AC143028 Macaca mu
30	25	8.8	626	11	BV064692 S212P6014
31	19	6.7	58539	2	AC137903 Mus muscu
32	16	5.6	1377	3	AY060959 Drosophil
33	15	5.3	858	6	BD251987 Screen me
34	15	5.3	858	6	AX032837 Sequence
35	15	5.3	1228	6	AX780306 Sequence
36	15	5.3	1263	9	BC001435 Homo sapi
37	15	5.3	1270	10	AF004106 Mus muscu
38	15	5.3	1310	10	BC003328 Mus muscu
39	15	5.3	1342	9	AK098671 Homo sapi
40	15	5.3	1351	9	AX779950 Sequence
41	15	5.3	1351	9	AF087894 Homo sapi
42	15	5.3	1376	6	AX011712 Sequence
43	15	5.3	1376	6	BD226323 Pancrati
44	15	5.3	1397	6	AR379492 Sequence
45	15	5.3	1709	9	AK026191 Homo sapi

ALIGNMENTS

RESULT 1

VLINKKVDs"

ORIGIN

Alignment Scores:

Pred. No.: 1,25e-306 Length: 858
 Score: 285.00 Matches: 285
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-889-733B-2 (1-285) x AX032835 (1-858)

QY 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
 DB 1 ATGGCGGGCTGGCGGACCCCTCCGCTTCCGCGGGGCCACCCAGCCCGCTGGCGGGCG 60
 QY 21 LeuProGluSerLeuGlyHisAlaLeuArgSerAlaGlyGlyGluValValVal 40
 DB 61 CTACCCGAGTCGCTCTGCCAGCAGCGCTGAGAGGCCCAAGCGCGAGGAGTGGACGTC 120
 QY 41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValValLeuGlySerLysLeuGlyLeu 60
 DB 121 GCCCGCGGGAACGGCAGCAGCAGCTCTACGTGGCGGTCTGGCGCAGCAAGCTGGCGCTG 180
 QY 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
 DB 181 CAGGTGGTGGAGTGGCGGCGGACGAGGCTTCCGACTGGCTCTTCCTGGAGGACGTG 240
 QY 81 AlaValValCysGluGluThrAlaLeuThrArgProGluValProSerArgArgLys 100
 DB 241 GCGTGGTGTGGAGGAGCGGCGCTCATCCCGCGCGGCGCGGCGGAGGAG 300
 QY 101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
 DB 301 GAGGTTCACATGATGAAGAAGCATTAGAAAACTTCAGTCAATATAGTAGAGATGAAA 360
 QY 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheVal 140
 DB 361 GATGAAATTCGAATCTAGATGGCGGAGATGTTTATTCAGCGCAGAGAAATTTTGTG 420
 QY 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuAlaAspThrPheLysAsp 160
 DB 421 GGCTTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGCTGATACTTTTAAGGAC 480
 QY 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
 DB 481 TATGCACTCTCCAGTGCAGCGCCAGATGGGTGGTTCATTTGAAGAGTTCTGCAGCATG 540
 QY 181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200
 DB 541 GCTGGGCTTAACCTGATCGCAATGGCTTAGTGAATCTGCACAGAAGGCCCTTAAGATC 600
 QY 201 MetGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAla 220
 DB 601 ATGCACAGATGATGATGACCCCGCTACGAAACTCACTGCTGCTGATGATGATGATGATG 660
 QY 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
 DB 661 AACTGTTATATCTTAATATCCCAACAAAGGCGACGCTTGTGTGACCGAACCCCGGAA 720
 QY 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal 260
 DB 721 GAGTATCCAGAAAGTGCAAGGTTTATGAGAAACTGAAGGACCATATGCTGATCCCGCTG 780
 QY 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuIleAsn 280
 DB 781 AGCATGTCTGAACGTGAAAGGTGATGGGTGCTGCTACCTGCTGCTGCTGCTGCTGCTGCT 840
 QY 281 LysLysValAspSer 285
 DB 841 AAGAGGTGAGTACTCC 955

RESULT 3
 BC033680
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BC033680 1624 bp mRNA linear PRI 07-OCT-2003
 Homo sapiens dimethylarginine dimethylaminohydrolase 1, mRNA (cdna
 clone MGC:45161 IMAGE:5189970), complete cds.
 BC033680
 BC033680.1 GI:21707414
 MGC
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heieh, F.,
 Dickson, K.L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Tadin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullen, S.O., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Prepared by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Akter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghini, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsurguen, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAP Plate: 68 Row: m Column: 1
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6912327.

FEATURES

source

1..1624
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"


```

/clone="MGC:45161 IMAGE:5189970"
/tissue type="Colon, Kidney, Stomach, adult, whole pooled"
/clone_lib="NIH_MGC_116"
/lab_host="DH103"
/notes="Vector: PCMV-SPORT6"
1..1624
/gene="DDAH1"
/notes="synonym: DDAH"
/db_xref="locusID:23576"
/db_xref="MIM:604743"
100..957
/codon_start=1
/product="DDAH1 protein"
/protein_id="AAH33680.1"
/db_xref="GI:21707415"
/db_xref="LocusID:23576"
/translation="WAGLGHFAAFGRATHAVVRALPESLIGQALRSKGEVDVARAE
RHOQLVVGIGSKLGVVLEPADSLPCVFDVAVVCEETALITRPGASRRKEV
DMWKEALKLQINIVEMDENATLDGGDLFTGREFFVLGSKRTNORGABILADTFKD
YANSTVPADGGLHLHSFSGMAGPNLIAIGSSSEKAKALKIMQMSDHEYDKLTVPDPI
ANCIYLINPKNGHVLLHRTPEYFESAKVYKLRKDHMLIIPVSMSELEKVDGLLTCCS
VLINKKVD"
133..936
/misc_feature
133..936
/notes="Aminotransferase; Region: Aminotransferase. This
family contains glycine (EC:2.1.4.1) and inosamine
(EC:2.1.4.2) aminotransferases, enzymes involved in
creatine and streptomycin biosynthesis respectively. This
family also includes arginine deiminases, EC:3.5.3.6.
These enzymes catalyse the reaction: arginine + H2O <=>
citrulline + NH3. Also found in this family is the
Streptococcus anti tumor glycoprotein"
/db_xref="CDD:pfam02274"

ORIGIN
Alignment Scores:
Pred. No.: 1.83e-277 Length: 1624
Score: 259.00 Matches: 259
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 90.88% Indels: 0
DB: 9 Gaps: 0

US-09-889-733B-2 (1-285) x BC033680 (1-1624)

QY 27 GlnHisAlaLeuArgSerAlaLysGlyGluValaspValAlaArgAlaGluArgGln 46
DB 178 CAGCAGCGCTGAGAGCGCCAGGCGGAGGAGTGAGCTGCCCGCGGAGCGGAG 237
QY 47 HisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuPro 66
DB 238 CACCAGCTCTACGTGGCGCTGTGGCGAGCAAGCTGGGCTGCAGGTGTGGAGCTGCCG 297
QY 67 AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGlu 86
DB 298 GCCAGCAGAGAGCTTCGGACTGCGTCTTCGTGGAGGAGCTGGCGGTGGTGGCAGGAG 357
QY 87 ThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLysGluValValAspMetLys 106
DB 358 ACGGCCCTCATCCCGAGCCCGGGCGCGGAGCGGAGGAGGAGGTGCACATGATGAAA 417
QY 107 GluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeu 126
DB 418 GAAGCATTAGAAAACCTTCAGCTCAATATAGTAGATGAAAGATGAAATCACTTTA 477
QY 127 AspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLysArgThr 146
DB 478 GATGGCGGAGATGTTTTATTACAGGCAGCAGAGATTTTTGTGGCGCTTCCAAAAGGACA 537
QY 147 AsnGlnArgGlyAlaGluLeuLeuAlaAspThrPheLysAspTyrAlaValSerThrVal 166
DB 538 AATCAACAGAGGCTGCTGAAATCTTGGCTGATCTTTTAAAGGACTATGCAGTCTCCACAGTG 597
QY 167 ProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIle 186

```

```

DB 598 CCAAGTGGCAGATGGTTCATTGGAAGAGTTCTTCAGCATGGCTGGGCTAACCTGATC 657
QY 187 AlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAsp 206
DB 658 GCAATTGGGTCTAGTGAATCTGCACAGAAGGCCCTTAAGATCATGCAACAGATGAGTGAC 717
QY 207 HisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsn 226
DB 718 CACCGGTACGACAAACTCACTGTGCTGATGACATAGCAGCAAACTGTATATATCTAAAT 777
QY 227 IleProAsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAla 246
DB 778 ATCCCAACAAAGGCGACGCTTCCTGCACCGAACCCCGGAGAGATATCCAGAAAGTGCA 837
QY 247 LysValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGlu 266
DB 838 AAGCTTTATGAGAACTGAAGGACCATATGCTGATCCCGTGAGCATGTGTGAAGTGGAA 897
QY 267 LysValAspGlyLeuLeuThrCysCysSerValLeuIleAsnLysLysValAspSer 285
DB 898 AAGTGGATGGGTGCTCACTGCTGCTCACTGCTGCTCACTGCTGCTCACTGCTGCTGCTCC 954

RESULT 4
BD094073
LOCUS BD094073 1633 bp DNA linear PAT 27-AUG-2002
DEFINITION Shear stress-responsive DNAs.
ACCESSION BD094073
VERSION BD094073.1 GI:22639661
KEYWORDS WO 0125427-A/34.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1633)
Najima,H., Yoshisue,H., Obayashi,M., Ota,T., Kawabata,A.,
Sakurada,K., Kuga,T., Sekine,S., Nakamura,Y. and Sugan,S.
Shear stress-responsive DNAs
Patent: WO 0125427-A 34 12-APR-2001;
KYCWA HAKKO KOGYO CO LTD.HIROSHI NOJIMA,HAJIME YOSHISUE, MASAYA
OBAYASHI, TOSHIO OTA,AYAKO KAWABATA,KAZUHIRO SAKURADA,TETSURO KUGA,
SUSUMU SEKINE, YUSUKE NAKAMURA,SUMIO SUGANO
OS Homo sapiens (human)
PN WO 0125427-A/34
PD 12-APR-2001
PF 02-OCT-2000 WO 2000JP006840
PR 01-OCT-1999 JP 99P 280976
PI HIROSHI NOJIMA,HAJIME YOSHISUE,MASAYA OBAYASHI,TOSHIO OTA, PI
AYAKO KAWABATA,
PI KAZUHIRO SAKURADA,TETSURO KUGA,SUSUMU SEKINE,YUSUKE NAKAMURA,
PI SUMIO SUGANO
PC C12N15/12,C07K14/435,C07K16/18,C12P21/02,C12Q1/68,A61K38/00,
PC A61K39/395,
PC A61K48/00,A61P9/10,G01N33/50,G01N33/53
CC
CDS
FH Key Location/Qualifiers
FT CDS (323)..(1177)
FEATURES
source
1..1633
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1.84e-277 Length: 1633
Score: 259.00 Matches: 259
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 90.88% Indels: 0
DB: 6 Gaps: 0

US-09-889-733B-2 (1-285) x BD094073 (1-1633)

```

QY	27	GlnHisAlaLeuArgSerAlaLysGlyGluValAspValAlaArgAlaGluArgGln	46
Db	401	CAGCACGCGCTGAGAGCGCAAGGCGGAGAGGTGGACGTGCCCGCGGAGCGGAG	460
QY	47	HiseGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuPro	66
Db	461	CACCAGCTCTACGTGGGCGTGGGCGAGCAAGCTGGGCGTCAGGTGGTGGAGCTGCCG	520
QY	67	AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGlu	86
Db	521	GCCGACGAGAGCCTTCGGACTCGCTTCGTGGAGAGCGTGGCCGTGGTGGAGGAG	580
QY	87	ThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLysGluValAlaAspMetLys	106
Db	581	ACGGCCCTCATACCCGACCCGCGCGAGCGGAGGAGAGGTTCATCATGATGAA	640
QY	107	GluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeu	126
Db	641	GAAGCATTTAGAAAACCTTCAGCTCAATATAGTAGAGATGAAGATGAAGATGCACTTTA	700
QY	127	AspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLysArgThr	146
Db	701	GATGCGGAGATGTTTATTTCACAGCGAGAGAAATTTTGTGGGCGCTTCCAAAAGGACA	760
QY	147	AsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrVal	166
Db	761	AATCAACGAGGTGCTGAAATCTTGCTCATACTTTTAAGGACTATGCAGTCTCCACAGTG	820
QY	167	ProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIle	186
Db	821	CCAGTGGCAGATGGGTTCATTTGAAGAGTTTCTGCAGCATGGCTGGGCGCTAACCTGATC	880
QY	187	AlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAsp	206
Db	881	GCAATTGGGTCTAGTGAATCTGCACAGAGCGCCCTTAAGATCATGCACAGATGAGTGAC	940
QY	207	HiseArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsn	226
Db	941	CACCGCTAGCAAACTCACTGTGCTGATGACATAGCAGCAAACTGTATATATCTAAAT	1000
QY	227	IleProAsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAla	246
Db	1001	ATCCCCAACAAAGGCGACGTCTTGCTGCACGACCCCGAGAGATATCCAGAAAGTGCA	1060
QY	247	LysValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGlu	266
Db	1061	AAGGTTTATGAGAACTGAAGACCATATGCTGATCCCGGTGACATGCTGAATCGAA	1120
QY	267	LysValAspGlyLeuLeuThrCysCysSerValLeuLeuLeuLysLysValAspSer	285
Db	1121	AAGGTGGATGGGCTGCTCACCTGCTGCTCAGTTTATTAACAAGAAAGTAGACTCC	1177
RESULT 5			
LOCUS	AB001915	1633 bp	mRNA linear PRI 21-SEP-2000
DEFINITION	Homo sapiens mRNA for NG,NG-dimethylarginine dimethylaminohydrolase, complete cds.		
ACCESSION	AB001915		
VERSION	AB001915.1	GI:4160665	
KEYWORDS	NG,NG-dimethylarginine dimethylaminohydrolase.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Kimoto,M., Miyatake,S., Sasagawa,T., Yamashita,H., Okita,M., Oka,T., Ogawa,T. and Tsuji,H.		
TITLE	Purification, cDNA cloning and expression of human NG,NG-dimethylarginine dimethylaminohydrolase		
JOURNAL	Eur. J. Biochem. 258 (2), 863-868 (1998)		
MEDLINE	95089653		
PUBMED	9874257		

REFERENCE 2 (bases 1 to 1633)			
AUTHORS Kimoto,M.			
TITLE Direct Submission			
JOURNAL Submitted (15-MAR-1997) Masumi Kimoto, Okayama Prefectural University, Faculty of Health and Welfare Science, Department of Nutritional Science, Kuboki 111, Soja, Okayama 719-1197, Japan (E-mail:kimoto@fhw.oka.-pu.ac.jp, Tel:0866-94-2158, Fax:0866-94-2158)			
FEATURES			
Location/Qualifiers			
source 1..1633			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/tissue_type="Kidney"			
CDS 323..1180			
/EC_number="3.5.3.18"			
/codon_start=1			
/product="NG,NG-dimethylarginine dimethylaminohydrolase"			
/protein_id="BAA37117.1"			
/db_xref="GI:4160666"			
/translation="MAGLGHFAAAGRATHAVVRALPESLGOHALRSKGEVDDVARAE RQHQLYGVLSKGLQVVELPADBSLPDCVFVEDVAVCEETALITRPGSPSRKEV DMKEALEKLQINIVEMKDNATLDGGDLFTGREFFVGLSKRNQRAEILATDFKD YAVSTVPADGLHLKSFCSMAGPNLIAIGSSESAQKALKINQMSDHRDKLTVPDDI AANCITYLNI PNKGHVLLHRTPEYPESAKVYEKLKDHMLIPVMSLEKVDLLTCCS VLINKKVDS"			
ORIGIN			
Alignment Scores:			
Pred. No.: 1.84e-277 Length: 1633			
Score: 259.00 Matches: 259			
Percent Similarity: 100.00% Conservative: 0			
Best Local Similarity: 100.00% Mismatches: 0			
Query Match: 90.88% Indels: 0			
DB: 9 Gaps: 0			
US-09-889-733B-2 (1-285) x AB001915 (1-1633)			
QY	27	GlnHisAlaLeuArgSerAlaLysGlyGluValAspValAlaArgAlaGluArgGln	46
Db	401	CAGCACGCGCTGAGAGCGCAAGGCGGAGAGGTGGACGTGCCCGCGGAGCGGAG	460
QY	47	HiseGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuPro	66
Db	461	CACCAGCTCTACGTGGGCGTGGGCGAGCAAGCTGGGCGTCAGGTGGTGGAGCTGCCG	520
QY	67	AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGlu	86
Db	521	GCCGACGAGAGCCTTCGGACTCGCTTCGTGGAGAGCGTGGCCGTGGTGGAGGAG	580
QY	87	ThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLysGluValAlaAspMetLys	106
Db	581	ACGGCCCTCATACCCGACCCGCGCGAGCGGAGGAGAGGTTCATCATGATGAA	640
QY	107	GluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeu	126
Db	641	GAAGCATTTAGAAAACCTTCAGCTCAATATAGTAGAGATGAAGATGAAGATGCACTTTA	700
QY	127	AspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLysArgThr	146
Db	701	GATGCGGAGATGTTTATTTCACAGCGAGAGAAATTTTGTGGGCGCTTCCAAAAGGACA	760
QY	147	AsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrVal	166
Db	761	AATCAACGAGGTGCTGAAATCTTGCTCATACTTTTAAGGACTATGCAGTCTCCACAGTG	820
QY	167	ProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIle	186
Db	821	CCAGTGGCAGATGGGTTCATTTGAAGAGTTTCTGCAGCATGGCTGGGCGCTAACCTGATC	880
QY	187	AlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAsp	206
Db	881	GCAATTGGGTCTAGTGAATCTGCACAGAGCGCCCTTAAGATCATGCACAGATGAGTGAC	940

QY 207 HisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsn 226
 Db 941 CACCCTACGACAACTCAGTGTGCTGATGACATAGCAGCAAACTGTATATATCTTAAT 1000
 QY 227 IleProAsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAla 246
 Db 1001 ATCCCCCAACAAGGCGACGCTTGTGTGACCGAAGAGAGATATCCAGAAAGTGCA 1060
 QY 247 LysValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGlu 266
 Db 1061 AAGTTTATGAGAACTGAGGACCAATATGCTATCCCGGTGAGCATGCTCGAACTGGAA 1120
 QY 267 LysValAspGlyLeuLeuThrCysSerValLeuLeuAsnLysLysValAspSer 285
 Db 1121 AAGGTGGATGGGCTGCTCACCTGTGCTCAGTTTAAATTAACAAGAAAGTAGACTCC 1177

RESULT 6
 AX548046
 LOCUS
 DEFINITION Sequence 22 from Patent WO02066654.
 ACCESSION AX548046
 VERSION AX548046.1 GI:25813142
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 REFERENCE
 AUTHORS Astromoff, A., Au-Young, J., Baughn, M.R., Ding, L., Duggan, B.M.,
 Forsythe, I.J., Gietzen, K.J., Griffin, J.A., Lee, E.A., Lu, Y.,
 Richardson, T.W., Ring, H.Z., Sanjanwala, M.M., Swarnakar, A.,
 Walla, N.K., Warren, B.A., Xu, Y., Yue, H. and Zebardjian, I.
 Drug metabolizing enzymes
 Patent: WO 02066654-A 22 29-AUG-2002;
 Incyte Genomics, Inc. (US)
 TITLE
 JOURNAL
 FEATURES
 source
 1. 4208
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 7671089CB1"

ORIGIN
 Alignment Scores:
 Pred. No.: 4,47e-277 Length: 4208
 Score: 259.00 Matches: 259
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 90.88% Indels: 0
 DB: 6 Gaps: 0

US-09-889-733B-2 (1-285) x AX548046 (1-4208)

QY 27 GlnHisAlaLeuArgSerAlaLysGlyGluValAspValAlaArgAlaGluArgGln 46
 Db 446 CAGCAGCGCTGAGAGGCCAGCGGAGGAGTGACGTGCGCGCGGAGACGCGAG 505
 QY 47 HisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuPro 66
 Db 506 CACAGCTCTACGTGGGGCTGCTGGGACAGCTGGGGCTGCGAGGTGGAGCTCCG 565
 QY 67 AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValCysGluGlu 86
 Db 566 GCCCAGCAGAGCCCTCCGAGCTGCGTCTTCGTGAGGACGTGGCGGTGTCGAGGAG 625
 QY 87 ThrAlaLeuIleThrArgProGlyValAspSerArgGlyValValAspMetMetLys 106
 Db 626 ACGCCCTTCATACCCGACCCCGGCGCGGAGGAGGAGGTGTGATCATGATGAA 685
 QY 107 GluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeu 126
 Db 686 GAAGCATTTAGAAAATCTCAGCTCAATATAGTAGAGATGAAGATGAATGAACCTTTA 745

QY 127 AspGlyGlyAspValLeuPheThrGlyArgGluPhePheValClyLeuSerLysArgThr 146
 Db 746 GATGGCGGAGATGTTTATTTCAGCAGCAGAGAAATTTTGTGGGCTTTCCAAAAGACA 805
 QY 147 AsnGlnArgGlyValAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrVal 166
 Db 806 AATCAACGAGGTGCTGAAATCTGGCTGATACCTTTTAAGGACTATGTCAGTCTCCACAGTG 865
 QY 167 ProValAlaAspGlyLeuHisLysSerPheCysSerMetAlaClyProLeuLeu 186
 Db 866 CCAAGTGGCAGATGGGTTCATTGAAGAGTTTCTGCAGCATGGCTGGGCTAACCTGATC 925
 QY 187 AlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAsp 206
 Db 926 GCAATTTGGGTAGTGAATCTGCACAGAGAGGCCCTTAAGATCATGCAACAGATGAGTGAC 985
 QY 207 HisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsn 226
 Db 986 CACCGCTACGACAACTCACTGTGCTGATGATGACAGCAAACTGTATATATCTTAAT 1045
 QY 227 IleProAsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAla 246
 Db 1045 ATCCCCCAACAAGGCGACGCTTGTGTGACCGAAGAGAGTATCCAGAAAGTGCA 1105
 QY 247 LysValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGlu 266
 Db 1106 AAGTTTATGAGAACTGAAGACCATATGCTGATGATCCCGTGAGCATGTCTGAACCTGAA 1165
 QY 267 LysValAspGlyLeuLeuThrCysCysSerValLeuLeuAsnLysValAspSer 285
 Db 1166 AAGTTGGATGGGCTGCTCACCTGTGCTCAGTTTAAATTAACAAGAAAGTAGACTCC 1222

RESULT 7
 BC043235
 LOCUS
 DEFINITION Homo sapiens dimethylarginine dimethylaminohydrolase 1, mRNA (cdna
 clone IMAGE:5295723), partial cds.
 ACCESSION BC043235
 VERSION BC043235.2 GI:34783628
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 3737)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullaly, S.J., Bosak, S.A., McSwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahy, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouford, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, J., Smalusz, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 MEDLINE
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 3737)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (09-JAN-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Sep 16, 2003 this sequence version replaced gi:28175756.
 Contact: MGC help desk
 Email: cgabs@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Teai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRK plate: 76 Row: j Column: 15.

FEATURES

Location/Qualifiers
 1..3737
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5295723"
 /tissue_type="Testis"
 /clone_lib="NIH_MGC_97"
 /lab_host="DH10B"
 /note="Vector: pBluescript"
 1..3737
 /gene="DDAH1"
 /note="synonym: DDAH"
 /db_xref="LocusID:23576"
 /db_xref="MIM:604743"
 1..741
 /gene="DDAH1"
 /codon_start=1
 /product="DDAH1 protein"
 /protein_id="AAH43235.2"
 /db_xref="GI:34783629"
 /db_xref="LocusID:23576"
 /db_xref="MIM:604743"
 /translation="VARAERQHLVYGLVSLGLQVVELPADESLPCVFVEDVAVV
 CEETALITRPAPSRKKEVDMMKALEKLQINIVEMKDENATLDGVDLFTGREFFVG
 LSKRTNORGAETLADTFDVAIVTPVADGLHLKSFCSMAGPNLIAIGSSSAOKALK
 IMQMSDHRKYDKLTPVDDIANCIYLNIPKGVLLHRTPEYFESAKVVEKLDHML
 IPVSNSELEKVDGLLTCCSVLINKVDS"
 1..720
 /gene="DDAH1"
 /note="Aminotransf; Region: Aminotransferase. This family contains glycine (EC:2.1.4.1) and inosamine (EC:2.1.4.2) aminotransferases, enzymes involved in creatine and streptomycin biosynthesis respectively. This family also includes arginine deiminases, EC:3.5.3.6. These enzymes catalyse the reaction: arginine + H2O <=> citrulline + NH3. Also found in this family is the Streptococcus anti tumor glycoprotein"
 /db_xref="CDD:pfam02274"

misc_feature

ORIGIN

Alignment Scores:
 Pred.No.: 1,148-262 Length: 3737
 Score: 246.00 Matches: 246
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 86.32% Indels: 0

DB: 9 Gaps: 0

US-09-889-733B-2 (1-285) x BC043235 (1-3737)

QY 40 ValAlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGly 59
 Db 1 GTCCCCCGCCGGAACGCGACACCGCTCTAGTGGCGCTGCTGGCGACCAAGCTGGG 60
 QY 60 LeuGlnValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAsp 79
 Db 61 CTGCAGGTGGTGGAGCTCCCGCGCAGCAGAGCTCCCGACTCGCTTCGTGGAGGAC 120
 QY 80 ValAlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArg 99
 Db 121 GTGGCCCTGTGTGGAGGAGAGCGGCCCTCATCCCGACCCCGGGCGCGAGCGAGG 180
 QY 100 LysGluValAspMetLysGluAlaLeuGluLysLeuGlnLeuAsnLeuValGluMet 119
 Db 181 AAGGAGGTTGACATGATGAAGAAGCATTAGAAAACCTTCAGCTCAATATAGTAGAGATG 240
 QY 120 LysAspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePhe 139
 Db 241 AAAGATGAAATGCACTTTAGATGGCGGAGATGTTTATTTCACAGCAGAGAAATTTT 300
 QY 140 ValGlyLeuSerLysArgThrAsnGlnAtGlyValGluLeuAlaAspThrPheLys 159
 Db 301 GTGGCCCTTCCAAAAGGACAAATCAACAGGTGCTGAAATCTTGGCTGATATCTTTAAG 360
 QY 160 AspTyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSer 179
 Db 361 GACTATGCACTCCACAGTCCAGTGGCAGTGGGTTGCTTGAAGATTTCTGCAGC 420
 QY 180 MetAlaGlyProAsnLeuLeuAlaGlySerSerGluSerAlaGlnLysAlaLeuLys 199
 Db 421 ATGCTGGGCTTAACCTGATCGCAATTTGGTCTAGTGAATCTGCACAGAGGCCCTTAAG 480
 QY 200 IleMetGlnGlnMetSerAspHisA:GlyTyrAspLysLeuThrValProAspAspIleAla 219
 Db 481 ATCATGCAACAGATGAGTGACCAACCGCTACGCAAACTCACTGCTGCTGATGACATAGCA 540
 QY 220 AlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrPro 239
 Db 541 GCAACTGTATATCTAAATATCCCAACAAAGGCGAGCTGTGTGTGACCAACCCCG 600
 QY 240 GluGluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIlePro 259
 Db 601 GAAGAGTATCCAGAAAGTGCAAGGTTTATGAGAACTGAAGGACCATATGCTGATCCCC 660
 QY 260 ValSerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuIle 279
 Db 661 GTGAGCATGTCTGAACCTGGAAGAGTGGATGGCTGCTCACCTGCTGCTCAGTTTAAT 720
 QY 280 AsnLysLysValAspSer 285
 Db 721 AACAGAAAGTAGACTCC 738

RESULT 8

HSM808292

LOCUS

DEFINITION

EX648145

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

CONSRMT

TITLE

HSM808292 3908 bp mRNA linear PRI 30-AUG-2003
 Homo sapiens mRNA; cDNA DKFZp686N2176 (from clone DKFZp686N2176).

EX648145.1 GI:34367304

Homo sapiens (human)

Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 3908)

Ansoerge,W., Krieger,S., Regiert,T., Rittmuller,C., Schwager,B.,

Wiemann,S.

The German Human cDNA Consortium

Direct Submission

JOURNAL Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuerberg, GERMANY

COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;
Sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project

This clone (DKFZp686N176) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cDNA/>.

FEATURES
source
1..3908
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686N176"
/tissue_type="human endometrium carcinoma cell line"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"

ORIGIN

Alignment Scores:
Pred. No.: 1.02e-193 Length: 3908
Score: 184.00 Matches: 184
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 64.56% Indels: 0
DB: 9 Gaps: 0

US-09-889-733B-2 (1-285) x HSM808292 (1-3908)

QY 102 ValAspMetLeuValGluAlaLeuGluValLeuGlnLeuAsnIleValGluMetLeuAsp 121
Db 384 GTTGACATGATGAAGAAGCAATTAAGAACTTCAGCTCAATATAGTAGAGATGAAGAT 443

QY 122 GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheValGly 141
Db 444 GAAATGCACTTTAGATGGCGAGATGTTTATTTCACAGCAGAGAAATTTTGTGGGC 503

QY 142 LeuSerIysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyr 161
Db 504 CTTTCCAAAGGACAAATCAACAGGGTGCTGAAATCTTGCTGATCTTTTAAGGACTAT 563

QY 162 AlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAla 181
Db 564 GCAGTCTCCACAGTGCAGTGGCAGATGGGTGCTTGAATTAAGAGTTCTCAGCATGGCT 623

QY 182 GlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMet 201
Db 624 GGGCTTAACCTGATCGCAATTGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATCATG 683

QY 202 GlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsn 221
Db 684 CAACAGATGATGACCCCGGTACCAACTCACTGCTGCTGATGATGATGATGATGATGAT 743

QY 222 CysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGluGlu 241
Db 744 TGTATATATCTAATATATCCCAACAAAGGCGACGCTTGTGTCACCGAACCCTCGAAGAG 803

QY 242 TyrProGluSerAlaLysValTyrGluLysLysAspHisValLeuIleProValSer 261
Db 804 TATCCAGAAAGTCAAAAGGTTTATGAGAACTGAGGACCATATGCTGATCCCGTGAGC 863

QY 262 MetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuIleAsnLys 281
Db 864 ATGTCGAACTGGAAGAGTGGTGGCTGCTCACCTGCTGCTCAGTTTAAATTAACAAG 923

QY 282 LysValAspSer 285
Db 924 AAGTAGACTCC 935

RESULT 9

BC034505 1120 bp mRNA linear ROD 25-NOV-2003
Mus musculus dimethylarginine dimethylaminohydrolase 1, mRNA (cDNA
clone MGC:30267 IMAGE:5134831), complete cds.

BC034505
BC034505.1 GI:21961623
MGC.
Mus musculus (house mouse)
Mus musculus
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1120)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, P.S., Wagner, J., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A.C., Rodrigues, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

TITLE
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932

JOURNAL
MEDLINE
PUBMED

REFERENCE
2 (bases 1 to 1120)
Strausberg, R.
Direct Submission
Submitted (24-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgs.nci.nih.gov>
Contact: MGC help Desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

FEATURES

source
1..1120
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:30267 IMAGE:5134831"
/tissue_type="Liver, normal. 5 month old male mouse."

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 41 Row: 0 Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.

```

/clone_lib="NCI_CGAP_L19"
/lab_host="DHIOB"
/notes="Vector: pCMV-SPORT6"
1. 1120
/seqname="pDahl"
/notes="synonym: 2510015N06Rik"
/db_xref="LocusID:69219"
/db_xref="LocusID:1916469"
122. 979
/codon_start=1
/product="dimethylarginine dimethylaminohydrolase 1"
/protein_id="AAH34505.1"
/db_xref="GI:21961624"
/db_xref="LocusID:69219"
/translation="MAGLHPSAFGRATHAVVRAPPESLCHRALRRSGBEVDFARAE
RQHELYVGLGSLGVQVPADESPLDVFEDVAVVCEETALITRPGAPSRKEV
DMKKEALQIQLNIVEMDENATLGGDLFTGREFFVGLSKRTNQRGAELTADTFD
YAVSTVPVADSLHLKSPCSMAGPNLIAIGSSSEAKALKIMQMSDRYDKLITVPDDM
RANCIVLNIPSKGHVLLHRTPEEPESAKVEKUKDHLILIPVNSEMEKVDGLLTCCS
VFINKKIDS"
misc_feature
155. 958
/notes="Amidinotransf; Region: Amidinotransferase. This
family contains glycine (EC:2.1.4.1) and inosamine
(EC:2.1.4.2) amidinotransferases, enzymes involved in
creatine and streptomycin biosynthesis respectively. This
family also includes arginine deiminases EC:3.5.3.6.
These enzymes catalyse the reaction: arginine + H2O <=>
citrulline + NH3. Also found in this family is the
Streptococcus anti tumour glycoprotein"
/db_xref="CDD:pfam02274"

ORIGIN
Alignment Scores:
Pred. No.: 1,69e-107 Length: 1120
Score: 106.00 Matches: 152
Percent Similarity: 98.70% Conservative: 0
Best Local Similarity: 98.70% Mismatches: 1
Query Match: 37.19% Indels: 2
DB: 10 Gaps: 0

US-09-889-733B-2 (1-285) x BC034505 (1-1120)
QY 65 LeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAlaValCys 84
DB 314 CTGCGCGCGAGAGAGCGCTGCGCGACTGCGTGTTCGTGGAGAGCGTGGCGCGTGTGC 373
QY 85 GluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgGlyGluValAspMet 104
DB 374 GAGGAGAGCGCGCTATCATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 433
QY 105 MetLysGluAlaLeuGluLysLeuGlnLeuAsnLeuValGluMetLysAspGluAsnAla 124
DB 434 ATGAAGAGAGCGCTTGGAAAACCTTCAGCTCAACATAGTAGAGATGAAGATGAAGATGA 493
QY 125 ThrLeuAspGlyGlyAspValLeuPheThrGlyArgPhePheValGlyLeuSerLys 144
DB 494 ACITTTGGATGGTGGGAGCGCTTATTACAGAGGAGAGAGATTTTTCGCGCGCGCGCG 553
QY 145 ArgThrAsnGlnArgGlyAlaGluLeuAlaAspThrPheLysAspTyrAlaValSer 164
DB 554 AGAACAAATCAACAGAGTGTGAATCTTGCTGTGATCTTTTAAGAGCTACGAGCTCTCT 613
QY 165 ThrValProValAlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProAs 184
DB 614 ACAGTCCCTGTGGCGGATTC-TTTGCAATTTAAAGAGTTTCTGAGAGATGGCGGAGCCAA 672
QY 184 nLeuLeuAlaLeuGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMe 204
DB 673 CCTGATTCGAATAGGGTCCAGCGAATCTGCAGAGAGCGCGCTCAGATCATGCAACAGAT 732
QY 204 tSerAspHisArgTyrAspLysLeuThrValProAspAsp 217
DB 733 GAGTGACCACTGTTATGACAAGCTCACTGTACCGGAGC 772

```

```

RESULT 10
AX683109 3008 bp DNA linear PAT 29-MAR-2003
LOCUS Sequence 83 from Patent EP1279744.
DEFINITION AX683109
ACCESSION AX683109
VERSION AX683109.1 GI:29370146
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1
AUTHORS Brooksbank R.A., Dixon A.K., Lee K. and Pinnock R.D.
TITLE Identification and use of molecules implicated in pain
JOURNAL Patent: EP 1279744-A 83 29-JAN-2003;
WARNER-LAMBERT COMPANY (US)
FEATURES
Location/Qualifiers
1..3008
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
/notes="N-G, N-G dimethylarginine dimethylaminohydrolase"

ORIGIN
Alignment Scores:
Pred. No.: 4,26e-107 Length: 3008
Score: 105.00 Matches: 152
Percent Similarity: 98.70% Conservative: 0
Best Local Similarity: 98.70% Mismatches: 1
Query Match: 37.19% Indels: 2
DB: 6 Gaps: 0

US-09-889-733B-2 (1-285) x AX683109 (1-3008)
QY 65 LeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAlaValCys 84
DB 624 CTGCGCGCGAGAGAGCGCTGCGTGTTCGTGGAGAGCGTGGCGCGTGTGC 683
QY 85 GluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLysGluValAspMet 104
DB 684 GAGGAGAGCGCGCTATCATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 743
QY 105 MetLysGluAlaLeuGluLysLeuGlnLeuAsnLeuValGluMetLysAspGluAsnAla 124
DB 744 ATGAAGAGAGCGCTTGGAAAACCTTCAGCTCAACATAGTAGAGATGAAGATGAAGATGA 803
QY 125 ThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLys 144
DB 804 ACCTTAGATGTGGGAGCGTCTTATTCAGCGAGAGAGTTTTCGCGCGCGCGCGCG 863
QY 145 ArgThrAsnGlnArgGlyAlaGluLeuAlaAspThrPheLysAspTyrAlaValSer 164
DB 864 AGGACAAATCAACAGAGTGTGAGATCTTGCTGCTATCTTCAAGGAGCTACGCGATTTCC 923
QY 165 ThrValProValAlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProAs 184
DB 924 ACAGTCCCGTGGCGGATTC-TTTGCAATTTAAAGAGTTTCTGAGAGATGGTGGCGGCG 982
QY 184 nLeuLeuAlaLeuGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMe 204
DB 983 CCTGATTCGAATAGGGTCCAGTGAATCTGCAGAGAGCGCGCTCAGATCATGCAACAGAT 1042
QY 204 tSerAspHisArgTyrAspLysLeuThrValProAspAsp 217
DB 1043 GAGTGACCACTGTTATGACAAGCTCACTGTACCGGAGC 1082

RESULT 11
D86041 3008 bp mRNA linear ROD 06-FEB-1999
LOCUS Rattus norvegicus mRNA for N-G,N-G-dimethylarginine
DEFINITION dimethylaminohydrolase, complete cds.

```

804 ACCTTAGATCGTGGGAGCGTCTCTATTCACGCGAGAGAGTTTTTTGTGGCCCTTTCCAAA 863

145 AtgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheIysAspTyrAlaValSer 164

864 AGACCAAAATCACAGAGGTGCTGAGATCTTGGCTGATCTTTCACAGGACTACGCAGATTCC 923

165 ThrValProValAlaAspGly-LeuHisLeuIysSerPheCysSerMetalGlyProAs 184

924 ACAGTCCCGCTGGCCGATTCTTCCTCATTTAAAGAGTTCTTCGACGACGCTGGCCCCAA 982

184 nLeuIleAlaIleGlySerSerGluSerAlaGlnIlyAlaLeuIysIleMetGlnGlnMe 204

983 CTTGATGCGAATAGGTCCAGTGAATCTGGCAGAGGCCCTCAAGATCATGCAACAGAT 1042

204 tSerAspHisArgTyrAspIysLeuThrValProAspAsp 217

1043 GAGTGACCCAGTTATGACAAGCTCACTGTACCGAGCAGC 1082

RESULT 12

AK098405

LOCUS AK098405 2895 bp mRNA linear PRI 12-SEP-2003

DEFINITION Homo sapiens cDNA FLJ25539 fis, clone CBR09481.

ACCESSION AK098405

VERSION AK098405.1 GI:21758410

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiya,M., Suzuki,Y., Hata,H., Nakagawa,K., Mizuno,S., Morinaga,M., Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A., Kawakami,B., Nagai,K., Isogai,T. and Sugano,S.

NEDO human cDNA sequencing project

Unpublished

2 (bases 1 to 2895)

Sugano,S. and Suzuki,Y.

Direct Submission

Submitted (08-JULY-2002) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:flcdnagene.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.

FEATURES

source

1..2895

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CBR09481"

/tissue_type="brain"

/clone_lib="CBR"

/note="Cloning vector: pMEISFL3"

ORIGIN

Alignment Scores:

Pred. No.:	1..19e-82	Length:	2895
Scores:	84.00	Matches:	97
Percent Similarity:	98.98%	Conservative:	0
Best Local Similarity:	98.98%	Mismatches:	0
Query Match:	29.47%	Indels:	1
DB:	9	Gaps:	0

rs-09-889-713B-2 (1-285) X AK098405 (1-2895)

QY 119 MetLysAspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGlu-Ph 138
DB 369 ATCAAGATCAAAATGCACTTTAGATGCGGAGATGTTTATTACAGGAGAGAAATT 428
QY 138 ePheValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPh 158
DB 429 TTTTGTGGGCTTCCAAAAGACAAATCAACAGAGGTGCTGAATCTTGCTGATACATT 488
QY 158 eLysAspTyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCy 178
DB 489 TAAGGACTATGAGTCTCCACAGTGCAGTGCAGATGGGTTGCATTTGAAGAGTTTCTG 548
QY 178 sSerMetAlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLe 198
DB 549 CAGCATGGTGGGCTTAACTGATCCATTTGGTCTAGTGAATCTGCACAGAGGCTT 608
QY 198 uLysIleMetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValPro 215
DB 609 TAAGATCATGCACAGATGAGTGACCCAGTACGACAACTCACTGTGCC 660

RESULT 13
AL162260/c
LOCUS
DEFINITION Homo sapiens chromosome 1 clone RP4-611J7 map p22.3-31.1, 8
unordered pieces.
ACCESSION AL162260
VERSION AL162260.4 GI:9863624
KEYWORDS HTG; HTGS_Phasel; HTGS_CANCELLED.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
McLay, K.
Direct Submission
Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Aug 21, 2000 this sequence version replaced gi:9212914.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: dj611J7
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: pGAP4; 100% of reads
Chemistry: Dye-terminator ET-amersham; 16% of reads Chemistry:
Dye-terminator Big Dye; 83% of reads
Consensus quality: 100691 bases at least Q40
Consensus quality: 101774 bases at least Q30
Consensus quality: 102472 bases at least Q20
Insert size: 103371; sum-of-contigs
Insert size: 112648; 4.2% error; agarose-fp
Quality coverage: 4.76x in Q20 bases; sum-of-contigs Quality
coverage: 4.41x in Q20 bases; agarose-fp

----- This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 4008: contig of 4008 bp in length
* 4009 4108: gap of 100 bp
* 4109 17620: contig of 13512 bp in length
* 17621 17720: gap of 100 bp
* 17721 29332: contig of 11612 bp in length

* 29333 29432: gap of 100 bp
* 29433 34595: contig of 5163 bp in length
* 34596 34635: gap of 100 bp
* 34636 51023: contig of 16328 bp in length
* 51024 51123: gap of 100 bp
* 51124 56471: contig of 5348 bp in length
* 56472 62357: contig of 100 bp
* 62358 62457: contig of 5786 bp in length
* 62458 104071: contig of 41614 bp in length.
FEATURES
source
1. 104071
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="p22.3-31.1"
/clone="RP4-611J7"
/clone_lib="RPC1-4"
1. 4008
/note="assembly fragment:00515"
fragment_chain:1
clone_end:SP6
vector_side:left
4109..17620
/note="assembly fragment:00727"
fragment_chain:1
17721..29332
/note="assembly fragment:00262"
fragment_chain:1
29433..34595
/note="assembly fragment:00257"
fragment_chain:1
34596..51023
/note="assembly fragment:00672"
fragment_chain:1
51124..56471
/note="assembly fragment:00205"
56572..62357
/note="assembly fragment:01056"
62458..104071
/note="assembly fragment:00375"
clone_end:T7
vector_side:right

ORIGIN

Alignment Scores:
Pred. No.: 2.66e-72 Length: 104071
Score: 76.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 26.67% Indels: 0
DB: 2 Gaps: 0
US-09-889-733B-2 (1-285) x AL162260 (1-104071)
QY 27 GlnHisAlaLeuArgSerAlaLysGlyGluValAspValAlaArgAlaGluArgGln 46
DB 37494 CAGCAGCGCTGAGAGCGCCAGGCGAGGTTGGAGCTGCCCGCGCGAAGCGGAG 37435
QY 47 HisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuPro 66
DB 37434 CACCAGCTTACGTGGCGGTGTGGGCGAGCAAGCTGGGGCTGCAGGTGGTGGAGTGGCG 37375
QY 67 AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValAlaValCysGluGlu 86
DB 37374 GCCGACGAGAGCTTCCCGACTGCTCTTCGTGGAGGACGTGGCGCGTGGTGTGGCAGGAG 37315
QY 87 ThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLysGluVal 102
DB 37314 ACGGCCCTCATCACCCGACCGGGCGCGCGAGCGCGAGAGGAGGTA 37267

RESULT 14


```

AC019261
LOCUS       AC019261             158405 bp    DNA    linear    HTG 12-MAR-2000
DEFINITION Homo sapiens clone RP11-2P7, WORKING DRAFT SEQUENCE, 38 unordered
            pieces.
ACCESSION   AC019261
VERSION     AC019261.3   GI:7229998
KEYWORDS    HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 158405)
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS     Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE       Homo sapiens, clone RP11-2P7
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 158405)
            Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavskiy,I., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lander,E., Lehotzky,J., Levine,R., Liu,C., Liu,G., Locke,K.,
            Macdonald,P., Margulis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
            Direct Submission
            Submitted (31-DEC-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            On Mar 12, 2000 this sequence version replaced gi:6899774.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ----- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIER
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            ----- Project Information
            Center project name: L2694
            Center clone name: 2_P_7
            ----- Summary Statistics
            Sequencing vector: M13; M77815; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.960731
            Consensus quality: 135380 bases at least Q40
            Consensus quality: 146343 bases at least Q30
            Consensus quality: 150850 bases at least Q20
            Insert size: 173000; agarose-fp
            Insert size: 154705; sum-of-contigs
            Quality coverage: 2.8 in Q20 bases; agarose-fp
            Quality coverage: 3.2 in Q20 bases; sum-of-contigs
            -----
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 38 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *
            * 1 1607: contig of 1607 bp in length
            * 1608 1707: gap of 100 bp
            * 1708 3028: contig of 1321 bp in length
            * 3029 3128: gap of 100 bp
            * 3129 4889: contig of 1761 bp in length
            *
            FEATURES
            source
            1..158405
            /organism="Homo sapiens"

```

```

4890 4989: gap of 100 bp
4990 6333: contig of 1444 bp in length
6334 8533: gap of 100 bp
8534 8085: contig of 1552 bp in length
8086 8185: gap of 100 bp
8186 9968: contig of 1783 bp in length
9969 10068: gap of 100 bp
10069 11857: contig of 1789 bp in length
11858 11957: gap of 100 bp
11958 14371: contig of 2414 bp in length
14372 14471: gap of 100 bp
14472 16047: contig of 1576 bp in length
16048 16147: gap of 100 bp
16148 18676: contig of 2529 bp in length
18677 18776: gap of 100 bp
18777 21785: contig of 3009 bp in length
21786 21885: gap of 100 bp
21886 24973: contig of 3088 bp in length
24974 25073: gap of 100 bp
25074 27441: contig of 2368 bp in length
27442 27541: gap of 100 bp
27542 30718: contig of 3177 bp in length
30719 30818: gap of 100 bp
30819 33922: contig of 3174 bp in length
33923 34092: gap of 100 bp
34093 36731: contig of 2639 bp in length
36732 39283: contig of 2452 bp in length
39284 41388: contig of 2455 bp in length
41389 44309: contig of 2271 bp in length
44310 44309: gap of 100 bp
44311 47829: gap of 100 bp
47830 50393: contig of 2564 bp in length
50394 50493: gap of 100 bp
50494 56065: contig of 5572 bp in length
56066 61306: contig of 5141 bp in length
61307 61406: gap of 100 bp
61407 65295: contig of 3889 bp in length
65296 71547: contig of 6152 bp in length
71548 71648: gap of 100 bp
71649 77272: contig of 5525 bp in length
77273 80143: contig of 2871 bp in length
80144 80243: gap of 100 bp
80244 84307: contig of 4064 bp in length
84308 84407: gap of 100 bp
84408 88430: contig of 4023 bp in length
88431 88530: gap of 100 bp
88531 94075: contig of 5545 bp in length
94076 94175: gap of 100 bp
94176 100399: contig of 6124 bp in length
100400 107046: contig of 6647 bp in length
107047 107146: gap of 100 bp
107147 114514: contig of 7368 bp in length
114515 122515: contig of 7901 bp in length
122516 122615: gap of 100 bp
122616 131415: contig of 8800 bp in length
131416 131515: gap of 100 bp
131516 138482: contig of 6987 bp in length
138483 138582: gap of 100 bp
138583 144615: contig of 6033 bp in length
144616 144715: gap of 100 bp
144716 158405: contig of 13690 bp in length.
            Location/Qualifiers
            1..158405
            /organism="Homo sapiens"

```


/note="Sequence confirmed by AC073061 sequenced by WUGSC"
27601..27792
/note="Sequence confirmed by AC019261 sequenced by WIBR"
147415..147467
/note="Sequence from overlapping clone BA290M5 (AL162256).
Assembly confirmed by restriction digest."
161180..161206
/note="Sequence from overlapping clone BA290M5 (AL162256).
Assembly confirmed by restriction digest."

ORIGIN

Alignment Scores:

Pred. No.:	4,19e-72	Length:	168953
Score:	76.00	Matches:	76
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	26.67%	Indels:	0
DB:	9	Gaps:	0

US-09-889-733B-2 (1-285) x AL360219 (1-168953)

QY	27	GlnHisAlaLeuArgSerAlaLysGlyGluValValAspValAlaArgAlaGluArgGln	46
Db	74848	CAGCACGGCTGAGAAAGCGCCAAAGGGCGAGAGGTGGACGTGCCCGCGCGGAACGGCAG	74789
QY	47	HisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuPro	65
Db	74788	CACCAGCTCTACGTGGGCTCTCTGGGCAGCAAGCTGGGGCTGCAGGTGGTGGAGCTGCCG	74729
QY	67	AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGlu	86
Db	74728	GCCGACGAGAGCCCTCCGGACTGGCTCTTCGTGGAGGACGTGGCCGTGGTGTGCGAGGAG	74669
QY	87	ThrAlaLeuIleThrArgProGlyValAlaProSerArgLysGluVal	102
Db	74668	ACGGCCCTCATCACCCGACCCCGGGCCCGGAGCGGAGGAGGTA	74621

Search completed: June 8, 2004, 02:16:41
Job time : 3367 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 7, 2004, 23:27:37 ; Search time 376 Seconds
(without alignments)
3220.044 Million cell updates/sec

Title: US-09-889-733B-2

Perfect score: 285

Sequence: 1 MAGLHPSAFGRATHAVRA.....EKVDGLLTCCSVLKNKVD 285

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6745144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xl
-Q/cgn2_1/USPTO.spool_p/US09889733/runat.07062004.083556.8342/app_query.fasta_1.455
-DB=N_Geneseq_29Jan04 -Qfmt=fastap -SUFFIX=olip2n.rng -MINWATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09889733 @CGN.1.1.708 @runat.07062004.083556.8342 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_29Jan04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002s: *
7: Geneseqn2003as: *
8: Geneseqn2003bs: *
9: Geneseqn2003cs: *
10: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	285	100.0	858	3 AAA47655	Aaa47655 Dimethyla
2	259	90.9	858	9 ADE71465	Ad71465 Human Thr
3	259	90.9	1633	4 AAH02907	Aah02907 Human she
4	259	90.9	1633	6 ABQ88102	Abq88102 Human ost
5	259	90.9	1633	7 ABQ88102	Abq88102 Human ost
6	259	90.9	4208	6 AAL41494	Aal41494 Drug meta
7	225	78.9	3778	9 ADE53981	Ad53981 Human pro
8	203	71.2	3651	3 AAF16077	Aaf16077 Human pro

9	198	69.5	858	9 ADE71463	Ad71463 Human Met
10	106	37.2	3008	9 ADB53014	Ad53014 Primary r
11	106	37.2	3008	9 ADB79843	Ad79843 Rat N-G,N
12	49	17.2	83946	6 ABQ88101	Abq88101 Human ost
13	38	13.3	6373	5 ABA14642	Ab14642 Human ner
14	38	13.3	6381	5 ABA14641	Ab14641 Human ner
15	27	9.5	420	3 AAC10265	Aac10265 Human sec
16	26	9.1	444	6 ABN93945	Abn93945 Mouse clo
17	20	7.0	60	6 ABN36282	Abn36282 Human spl
18	16	5.6	873	4 ABL03335	Ab103335 Drosophil
19	15	5.3	452	8 ACH27659	Ach27659 Human adu
20	15	5.3	472	8 ACH21905	Ach21905 Human adu
21	15	5.3	845	3 AAA93930	Aaa93930 Human mel
22	15	5.3	858	3 AAA47656	Aaa47656 Dimethyla
23	15	5.3	1228	3 AAD00677	Aad00677 Human Hyd
24	15	5.3	1376	2 AAZ52967	Aaz52967 Human pro
25	14	4.9	655	9 AAD34806	Ad34806 Mouse mit
26	11	3.9	239	5 AAF68469	Aaf68469 Human lun
27	11	3.9	239	6 ABK38380	Abk38380 cDNA enco
28	11	3.9	239	7 ACA10709	Aca10709 Human lun
29	11	3.9	239	7 ABX93660	Abx93660 Lung canc
30	11	3.9	239	10 ADE71675	Ad71675 Human lun
31	11	3.9	862	6 ABQ47888	Abq47888 Oligonuc
32	11	3.9	862	6 ABQ47889	Abq47889 Oligonuc
33	11	3.9	898	6 ABQ41319	Abq41319 Oligonuc
34	11	3.9	898	6 ABQ41318	Abq41318 Oligonuc
35	10	3.5	770	6 ABS76755	Ab576755 Frog embr
36	10	3.5	3223	4 AAK73698	Aak73698 Human imm
37	10	3.5	3643	4 ABL03334	Ab103334 Drosophil
38	10	3.5	28772	6 ABK83555	Abk83555 Human CDN
39	9	3.2	456	4 AAS55803	Aas55803 Streptoco
40	9	3.2	641	2 AAQ05515	Aaq05515 Upstream
41	9	3.2	664	2 AAQ05513	Aaq05513 Upstream
42	9	3.2	664	2 AAQ05514	Aaq05514 Upstream
43	9	3.2	664	2 AAV18975	Aav18975 Nucleotid
44	9	3.2	774	6 ABQ86223	Abq86223 Lactobaci
45	9	3.2	774	9 ADE12836	Ad12836 DNA encod

ALIGNMENTS

RESULT 1

AAA47655
ID AAA47655 standard; cDNA; 858 BP.

XX
AC AAA47655;

XX
DT 08-NOV-2000 (first entry)

XX
DE Dimethylarginine dimethylaminohydrolase (DDAH1) coding sequence.

XX
KW Dimethylarginine dimethylaminohydrolase; DDAH; DDAH1; DDAH2;
KW arginine deaminase; hyperlipidemia; renal failure; hypertension;
KW restenosis; atherosclerosis; schizophrenia; multiple sclerosis; cancer;
KW ischemia reperfusion injury; septic shock; multi organ failure;
KW arthritis; skin disorders; inflammatory cardiac disease; migraine;
XX infection; ds.

XX
OS Homo sapiens.

XX
FH Key Location/Qualifiers

XX
FT CDS 1..858

XX
FT /*tag= a

XX
FT /product= "DDAH1"

XX
XX WO2000044888-A2.

XX
XX 03-AUG-2000.

XX
XX 26-JAN-2000; 2000WO-GB000226.

XX
XX 26-JAN-1999; 99GB-00001705.

XX
XX 04-JUN-1999; 99GB-00013066.

XX PA (UNLO) UNIV COLLEGE LONDON.
 XX PI Vallance PUT, Leiper JM, Whitley GSJ, Charles IG;
 XX PI WPI: 2000-543392/49.
 DR P-PSDB; AAB01477.
 XX Novel methylarginase polypeptides and polynucleotides, used to identify
 PT modulators of them, which are used in the treatment of e.g. cancer,
 PT hypertension, and bacterial infections.
 XX Claim 1; Page 55-56; 68pp; English.
 XX Nucleotides encoding methylarginase polypeptides, vectors comprising
 CC these nucleotides and the polypeptides themselves can be used in
 CC medicaments for the treatment of hyperlipidemia, renal failure,
 CC hypertension, restenosis after angioplasty, atherosclerosis,
 CC complications of heart failure, schizophrenia, multiple sclerosis or
 CC cancer. Modulators of the enzyme can be used in medicaments for the
 CC treatment of ischemia-reperfusion injury of the brain or heart, cancer,
 CC lethal hypertension in severe inflammatory conditions such as septic
 CC shock or multi-organ failure, or local and systemic inflammatory
 CC disorders including arthritis, skin disorders, inflammatory cardiac
 CC disease, migraine, or microbial or bacterial infection
 XX SQ Sequence 858 BP; 217 A; 214 C; 256 G; 171 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3,71e-281 Length: 858
 Score: 285.00 Matches: 285
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-09-889-733B-2 (1-285) x AAA47655 (1-858)

QY 1 MetAlaGlyLeuGlyHisProSerAlaPheGlyArgAlaThrHisAlaValValArgAla 20
 DB 1 ATGGCGGGCTGGCCACCCCTCCGGCTTCGGCCGGGCCACCCACCGCCGCTGGTGGGGCG 60

QY 21 LeuProGluSerLeuCysGlnHisAlaLeuArgSerAlaLysGlyGluGluValAspVal 40
 DB 61 CTACCCGAGTCGCTCCGACGACGCGCTGAGAGCGGCCCAAGGGCGAGGAGGTGACGTC 120

QY 41 AlaArgAlaGluArgGlnHisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeu 60
 DB 121 GCCCGCGGACGGCAGCAGCAGCTCTACGTCGGCGCTGCTGGCAGCAAGCTGGGGCTG 180

QY 61 GlnValValGluLeuProAlaAspGluSerLeuProAspCysValPheValGluAspVal 80
 DB 181 CAGGTGGTGGAGTGGCGGCCGACGAGAGCCCTCCGGACTGGCTCTCGTGGAGGACGTG 240

QY 81 AlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLys 100
 DB 241 GCGTGGTGTGGAGGAGACGGCCCTCATCACCCGCGGGCGCCGAGCGCGGAGGAG 300

QY 101 GluValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
 DB 301 GAGGTGGATGATGATGAAGAGCATTAGAAAACCTCAGCTCAATATAGTAGAGATGAA 360

QY 121 AspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheVal 140
 DB 361 GATGAAATGCACCTTAGATGGCGAGATGTTTTATTACAGGCGAGAGAAATTTTGTG 420

QY 141 GlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuAlaAspThrPheLysAsp 160
 DB 421 GGCCTTTCCAAAGGACAATCAACGAGGTGCTGAAATCTTGCTGATACCTTTAAGGAC 480

QY 161 TyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMet 180
 DB 481 TATGAGTCTCCACAGTGGCAGTGGCAGATGGGTTCATTTGAAGATTTCTGCAGCATG 540

QY 181 AlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIle 200
 DB 541 GCTGGGCTTAACCTGATCGCAATTGGTCTAGTGAATCTGCACAGAAGCCCTTAAGATC 600

QY 201 MetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAlaIle 220
 DB 601 ATCCAAACATGATGATGACCCAGCCCTACGACAACCTCACTGTGCTGATGACATAGCAGCA 660

QY 221 AsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrProGlu 240
 DB 661 AACTGTATATATCTAAATATCCCAACAAGGCGACGTCTGCTGCACCAACCCCGGAA 720

QY 241 GluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHisMetLeuIleProVal 260
 DB 721 GAGTATCCGAAAGTCAAGGTTTATGAGAACTGAGGACCATATGCTGATCCCGTG 780

QY 261 SerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCysSerValLeuIleAsn 280
 DB 781 AGCATGCTGAACCTGGAAAAGGTGGATGGCTGCTCACTGCTGCTCAGTTTAAATTAAC 840

QY 281 LysLysValAspSer 285
 DB 841 AAGAAGGTAGACTCC 855

RESULT 2

ADE71465
 ID ADE71465 standard; DNA; 858 BP.

XX AC ADE71465;

XX DT 29-JAN-2004 (first entry)

XX DE Human Thr87 wild-type dimethylarginine dimethylaminohydrolase-1 DNA.

XX KW Human; dimethylarginine dimethylaminohydrolase-1; DD4H 1; chromosome 1;
 KW single nucleotide polymorphism; SNP; ADMA;
 KW asymmetrical dimethyl-arginine; nitric oxide; cardiovascular disease;
 KW coronary heart disease; cerebrovascular disease; hypertension; diabetes;
 KW susceptibility; genotyping; therapeutic targeting; antidiabetic;
 KW cardiant; cerebroprotective; hypotensive; gene therapy; gene; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 XX CDS 1..858

XX FT /*tag= a
 XX FT /product= "Wild-type (Thr87) DD4H 1"
 XX FT variation replace(260,T)

XX FT /*tag= b
 XX FT /standard_name= "Single nucleotide polymorphism (SNP)"
 XX FT /note= "Results in Thr (C variant) or Met (T variant) at
 XX FT position 87 of the protein"

XX PN W02003089638-A1.

XX PD 30-OCT-2003.

XX PF 11-APR-2003; 2003WO-FI000274.

XX PR 19-APR-2002; 2002US-00125456.

XX XX (JURI-) JURILAB LTD OY.

XX XX Valkonen V, Salonen JT, Pirskanen M, Tuomainen T, Laakso J;

XX PI Laaksonen R;

XX XX WPI: 2003-854121/79.

XX DR P-PSDB; ADE71466.

XX XX New nucleic acid encoding a variant dimethylarginine
 PT dimethylaminohydrolase-1 (DD4H-1) protein, useful for treating diabetes,
 PT and its vascular complications, e.g. coronary or cerebrovascular disease

PT or hypertension.
 PS Disclosure; SEQ ID NO 3; 37bp; English.
 XX
 CC The invention relates to a variant dimethylarginine
 CC dimethylaminohydrolase-1 (DDAH 1; ADE71463) and a nucleic acid encoding
 CC it (ADE71463). DDAHs regulate the metabolism of ADMA (asymmetric
 CC dimethyl-arginine), a naturally occurring nitric oxide synthase inhibitor
 CC and nitric oxide, which are relevant in disorders such as cardiovascular
 CC disease and diabetes. The DDAH 1 gene is located on chromosome 1, which
 CC has been implicated in susceptibility to diseases such as familial
 CC combined hyperlipidaemia, premature coronary artery disease (CAD), non
 CC insulin-dependent (type 2) diabetes mellitus, and diastolic hypertension.
 CC The variant DDAH 1 of the invention has Met at position 87, rather than
 CC Thr, which is caused by a C to T change in exon 1 at a position
 CC corresponding to base 260 in ADE71463 and ADE71465. Individuals hetero-
 CC or homozygous for the Thr87Met allele have an increased risk of
 CC developing cardiovascular diseases and diabetes. The invention also
 CC relates to a DDAH 1 cDNA capture probe; a method and kit for diagnosing a
 CC susceptibility to a cardiovascular disease and diabetes in an individual
 CC by determining the DDAH 1 genotype; a method of treating diabetes or
 CC vascular complications of diabetes by enhancing nitric oxide
 CC availability, production or concentration; a method of targeting
 CC treatment of cardiovascular disease and diabetes in a hypertensive
 CC patient by determining the DDAH 1 genotype and treating them with a drug
 CC that affects nitric oxide availability, production or metabolism; and a
 CC transgenic animal which carries a human variant DDAH 1 nucleic acid
 CC sequence. The nucleic acid molecules and polypeptides are useful for
 CC treating cardiovascular disease (e.g., coronary heart disease,
 CC cerebrovascular disease, and hypertension), and diabetes and its vascular
 CC complications. The methods are useful for determining whether a patient
 CC will benefit from treatment with a drug which affects nitric oxide
 CC availability, production or metabolism; a drug which reduces ADMA
 CC availability or concentration; or an agent which elevates DDAH
 CC availability or concentration (such as DDAH agonist). The methods are
 CC also useful for determining whether a patient will be at risk of adverse
 CC effects if DDAH antagonists are administered. The present sequence
 CC represents a nucleic acid encoding the "wild-type" (Thr87) DDAH 1.
 XX
 SQ Sequence 858 BP; 218 A; 213 C; 258 G; 169 T; 0 U; 0 Other;

Alignment Scores:
 Pred No.: 1.32e-254 Length: 858
 Score: 259.00 Matches: 259
 Percent Similarity: 100.00% Conservations: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 90.88% Indels: 0
 DB: 9 Gaps: 0

US-09-889-733B-2 (1-285) x ADE71465 (1-858)

QY 27 GlnHisAlaLeuArgSerAlaLysGlyGluValAspValAlaArgAlaGluArgGln 46
 DB 79 CAGCAGCGCTGGAAGCGCCCAAGCGGAGAGGTGGACGTGGCGCTGGAGGTGGAGTGGCG 138
 QY 47 HisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuPro 66
 DB 139 CACCAGCTCTACGTGGGGGTCTGGCGAGCAAGCTGGGGCTGGAGGTGGAGTGGCG 198
 QY 67 AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGlu 86
 DB 199 GCCCAGCAGAGCGCTTCCGACTGGCTCTTCGTGGAGGACGTGGCGGTGGTGGAGGAG 258
 QY 87 ThrAlaLeuLeuThrArgProGlyValAlaProSerArgArgLysGluValAspMetLys 106
 DB 259 ACGGCCCTCATCCCGACCCCGGGCGCGAGCGGAGGAGGTGGATGATGAAA 318
 QY 107 GluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeu 126
 DB 319 GAAGCATTAGAAAACCTTCAGTCAATATAGTAGATGAAAGATGAAATGAAATGCAACTTTA 378
 QY 127 AspGlyGlyAspValLeuPheThrClyArgGluPheValGlyLeuSerLysArgThr 146

DB 379 GATGGCGGAGATGTTTTATTTCACAGGCAGAGAAATTTTTTGTGGCCCTTCCAAAAGGACA 438
 QY 147 AsnGlnArgGlyAlaGluLeuAlaAspThrPheLysAspTyrAlaValSerThrVal 166
 DB 439 AATCAACGAGGTGCTGAATAATCTTGCTGATACTTTTAAAGGACTATGCAGTCTCCACATG 498
 QY 167 ProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIle 186
 DB 499 CCAGTGGCAGATGGGTTCATTTGAAGAGTTTCTGCAGCATGGCTGGCCCTTAACCTGATC 558
 QY 187 AlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAsp 206
 DB 559 GCAATTTGGTCTAGTGAATCTGCACAGAGGCCCTTAAATCATGCAACACATGATGATGAC 618
 QY 207 HisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsn 226
 DB 619 CACCGCTACGACAAACTCACCTGTGCTGATGACATAGCAGCAAACTGTATATATCTAAAT 678
 QY 227 IleProAsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAla 246
 DB 679 ATCCCAACAAAGGCGACGTCTTGCTGCACCGAACCCCGGAGAGTATCCAGAAAGTGCA 738
 QY 247 LysValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGlu 266
 DB 739 AAGGTTTATGAAACTGAGGACCATATGCTGATCCCGTGAGCATGCTGAACTGGAA 798
 QY 267 LysValAspGlyLeuLeuThrCysCysSerValLeuIleAsnLysLysValAspSer 285
 DB 799 AAGGTGGATGGGTCTGCCTGCTGCTCAGTTTTTAATTAACAAGAAAGTAGACTCC 855
 RESULT 3
 AAH02907
 ID AAH02907 standard; DNA; 1633 BP.
 XX
 AC AAH02907;
 XX
 DT 15-JUN-2001 (first entry)
 XX
 DE Human shear stress-response coding sequence SEQ ID NO: 67.
 XX
 KW Human; shear stress-response protein; vascular disease; arteriosclerosis;
 KW ds.
 XX
 OS Homo sapiens.
 XX
 PN WC000125427-A1.
 XX
 PD 12-APR-2001.
 XX
 PF 02-OCT-2000; 2000WO-JP006840.
 XX
 PR 01-OCT-1999; 99JP-00280976.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 PA (NOJII) NOJIMA H.
 XX
 PI Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;
 PI Kuga T, Sekine S, Nakamura Y, Sugano S;
 XX
 DR WPI; 2001-266308/27.
 DR P-PSDB; AAB90784.
 XX
 PT DNA sequences, proteins encoded by them and antibodies against them
 PT useful in diagnosis and treatment of vascular disease caused by
 PT arteriosclerosis.
 XX
 PS Claim 20; Page 404-407; 678pp; Japanese.
 CC
 CC The present invention provides the protein and coding sequences of a
 CC number of human shear stress response proteins. These are useful in the
 CC diagnosis, treatment and screening of vascular diseases caused by
 CC arteriosclerosis, including heart failure, post-PTCA restenosis and
 CC hypertension

```
XX SQ Sequence 1633 BP; 400 A; 406 C; 432 G; 393 T; 0 U; 2 Other;
Alignment Scores:
Pred. No.: 2,45e-254 Length: 1633
Score: 259.00 Matches: 259
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 90.88% Indels: 0
DB: 4 Gaps: 0

US-09-889-733B-2 (1-285) x AAH02907 (1-1633)
QY 27 GlnHisAlaLeuArgSerAlaLysGlyGluGluValAlaAspValAlaArgAlaGluArgGln 46
Db 401 CAGCAGCGGCTGAGAAGGCCCAAGGGCGAGAGTGGACGTCGCCCGCGCGGAACGGCAG 460
QY 47 HisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuPro 66
Db 461 CACCAGCTCTACCTGGGGTGTCTGGGAGCAGCTGGGGCTGCAGGTGGTGGACTGCCG 520
QY 67 AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGlu 86
Db 521 GCCGACGAGAGCGCTTCGAGCTGCTTCCTGAGAGGACGTGGCGCTGGTGTGGAGGAG 580
QY 87 ThrAlaLeuLeuLeuThrArgProGlyAlaProSerArgArgLysGluValAlaAspMetMetLys 106
Db 581 ACGGCCCTCATCACCAGCCCGGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 640
QY 107 GluAlaLeuGluLeuGlnLeuAsnLeuValGluMetLysAspGluAsnAlaThrLeu 126
Db 641 GAAGCATTTAGAAATCTTCAGCTCAATATAGTAGAGATGAAGATGAAGATGAAGATTTA 700
QY 127 AspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLysArgThr 146
Db 701 GATGCGGAGAGTGTATTTCACGCGCAGAGAAATTTTGTGGGCGCTTTCACAAAGGACA 760
QY 147 AsnGlnArgGlyAlaGluLeuLeuAlaAspThrPheLysAspTyrAlaValSerThrVal 166
Db 761 AATCACGAGGCTGCTGAATCTTGCTGATCTTTTAGGACTATGCACTCCACAGTG 820
QY 167 ProValAlaAspGlyLeuHisLysSerPheCysSerMetAlaGlyProAsnLeuLeu 186
Db 821 CCAGTGGCAGATGGGTTCATTTGAAGAGTTCCTGCAGCATGGCTGGGCGCTTAACCTGATC 880
QY 187 AlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAsp 206
Db 881 GCAATTGGGTCTAGTGAATCTGCACAGAGGCGCTTAAATCATGCAACAGATGAGTGAC 940
QY 207 HisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsn 226
Db 941 CACCGCTACGACAACTCACTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1000
QY 227 IleProAsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAla 246
Db 1001 ATCCCCCAACAAAGGGGACGCTGTGTGACCAAGGACCCCGGAAGATATCCAGAAAGTGA 1060
QY 247 LysValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGlu 266
Db 1061 AAGGTTTATGAGAACTGAAGACCATATGCTGATGATGATGATGATGATGATGATGATG 1120
QY 267 LysValAspGlyLeuLeuThrCysCysSerValLeuLeuAsnLysLysValAspSer 285
Db 1121 AAGGTGATGGGTGCTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1177

RESULT 4
ID ABQ88102 standard; cDNA; 1633 BP.
XX AC ABQ88102;
XX DT 18-SEP-2002 (first entry)
XX
```

DE Human osteoblast differentiation related cDNA SEQ ID NO 9.
XX Human; osteoblast; stem cell differentiation; bone tissue deposition;
KW osteoporosis; osteopathic; ss.
OS Homo sapiens.
XX WO200250301-A2.
PN 27-JUN-2002.
PD 18-DEC-2001; 2001WO-US048276.
PF 18-DEC-2000; 2000US-0255882P.
PR 24-APR-2001; 2001US-0285691P.
XX (GENE-) GENE LOGIC INC.
PA (PROC) PROCTER & GAMBLE CO.
XX Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
PI Mertz L;
XX WPI; 2002-557663/59.

Use of genes and their expression profiles associated with osteoblast differentiation for screening modulators bone formation, for diagnosing or treating e.g. osteoporosis, or as markers for the differentiation process.

Claim 1; SEQ ID NO 9; 78pp + Sequence Listing; English.

The invention relates to genes and their expression profiles are used for: (a) screening modulators of precursor stem cell differentiation into osteoblasts, or bone tissue deposition; (b) diagnosing abnormal deposition of bone tissue, abnormal rate of osteoblast formation or osteoporosis; or (c) treating or monitoring treatment of the conditions cited in (b); or monitoring the progression of bone tissue deposition. Specific conditions include postmenopausal osteoporosis, glucocorticoid osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-induced abnormalities in bone formation or bone loss, conditions that involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis), skeletal disease linked to breast cancer, mastocytosis, Panconi syndrome or fibrous dysplasia. The present sequence is that of an osteoblast differentiation associated cDNA marker of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 1633 BP; 400 A; 406 C; 432 G; 393 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 2,45e-254 Length: 1633
Score: 259.00 Matches: 259
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 90.88% Indels: 0
DB: 6 Gaps: 0

US-09-889-733B-2 (1-285) x ABQ88102 (1-1633)

QY 27 GlnHisAlaLeuArgSerAlaLysGlyGluGluValAlaAspValAlaArgAlaGluArgGln 46
Db 401 CAGCAGCGGCTGAGAAGGCCCAAGGGCGAGAGTGGACGTCGCCCGCGGAACGGCAG 460
QY 47 HisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuPro 66
Db 461 CACCAGCTCTACCTGGGGTGTCTGGGAGCAGCTGGGGCTGCAGGTGGTGGACTGCCG 520
QY 67 AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGlu 86
Db 521 GCCGACGAGAGCGCTTCGAGCTGCTTCCTGAGAGGACGTGGCGCTGGTGTGGAGGAG 580
QY 87 ThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLysGluValAlaAspMetMetLys 106

Db 581 ACAGCCCTCATACCCAGCCCGGGGCGCGAGCGGAGGAGGTTGACATGATGAA 640
QY 107 GluAlaLeuGluLysLeuGlnLeuAsnLeuValGluMetLysAspGluAsnAlaThrLeu 126
Db 641 GAAGCATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAGATGAAATCCACTTTA 700
QY 127 AspGlyGlyValLeuPheThrGlyArgGluPheValGlyLeuSerLysArgThr 146
Db 701 GATGGCGAGAGTATTTATTCAGACGACAGAGATTTTGTGGGCTTTCCAAAGGACA 760
QY 147 AsnGluArgGlyAlaGluLeuAlaLeuAspThrPheLysAspTyrAlaValSerThrVal 166
Db 761 AATCAACAGAGTCTGAATCTTTGGCTGATACCTTTAAGGACTATGACAGTCTCCACAGTG 820
QY 167 ProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuLeu 186
Db 821 CCAGTGGCAGATGGTGGCTTTGAAGAGATTTCTGCAGCATGGCTGGGCTTACCTGATC 880
QY 187 AlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAsp 206
Db 881 GCAATGGGTCTAGTGAATCTGCACAGAGGCCCTTAAGATCATGCAACAGATGAGTGAC 940
QY 207 HisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsn 226
Db 941 CACCGCTACGACAACTCACTGCTGCTGATGACATAGCAGCAACTGTATATCTAAAT 1000
QY 227 IleProAsnLysGlyHisValLeuLeuHisArgThrProGluTyrProGluSerAla 246
Db 1001 ATCCCCAACAAGGCGACGCTTCTGCTGCACGAGACCCCGGAGAGTATCCAGAAAGTGA 1060
QY 247 LysValTyrGluLysLeuLysAspHisMetLeuLeuProValSerMetSerGluLeuGlu 266
Db 1061 AAGTGTATGAGAACTGAAGACCATATGCTGATCCCGTGAGCATGTCTGAACCTGGA 1120
QY 267 LysValAspGlyLeuLeuThrCysCysSerValLeuLeuLeuAsnLysValAspSer 285
Db 1121 AAGTGTGATGGCTGCTCACCTGCTGCTCAGTTTAAATTAACAGAAAGTAGACTCC 1177

RESULT 5
ABZ35939
ID ABZ35939 standard; cDNA; 3932 BP.
XX
AC ABZ35939;
XX
DT 10-FEB-2003 (first entry)
XX
DE Human secretory polynucleotide SPTM SEQ ID NO 103.
XX
KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;
KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;
KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;
KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;
KW anti-inflammatory; immunosuppressive; neuroprotective; neurotropic;
KW neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;
KW antipsychiatric; antianaemic; anti-HIV; human immunodeficiency virus;
KW secretory polynucleotide; secretory protein; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200283876-A2.
XX
PD 24-OCT-2002.
XX
PF 27-MAR-2002; 2002WO-US009921.
XX
PR 29-MAR-2001; 2001US-0280067P.
PR 29-MAR-2001; 2001US-0280068P.
PR 16-MAY-2001; 2001US-0291280P.
PR 17-MAY-2001; 2001US-0291829P.
PR 17-MAY-2001; 2001US-0291849P.
PR 19-JUN-2001; 2001US-0299428P.
PR 20-JUN-2001; 2001US-0299776P.

20-JUN-2001; 2001US-0300001P.
(INCY-) INCYTE GENOMICS INC.
XX
PA Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshay SR;
PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstein EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX
DR MPI; 2003-075543/07.
DR P-PSDB; ABP75492.
XX
PT New human secretory proteins and polynucleotides, useful for diagnosing,
PT treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),
PT neurological disorders (e.g. Alzheimer's), or cell proliferations or
PT cancers.
XX
PS Claim 1; SEQ ID NO 103; 458pp + Sequence Listing; English.
XX
CC The invention relates to a secretory polynucleotide (designated sptm)
CC comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a
CC naturally occurring polynucleotide sequence at least 90 % identical to
CC the polynucleotide sequence, a polynucleotide complementary to them or an
CC RNA equivalent of them. The polypeptide or polynucleotide are useful for
CC treating, preventing or diagnosing a disease or condition associated with
CC the expression of functional SPTM. These are particularly useful for
CC diagnosing, treating or preventing autoimmune/inflammatory disorders
CC (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's
CC disease), neurological disorders (e.g. epilepsy, Huntington's disease,
CC dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,
CC multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,
CC psoriasis, polycythemia vera, or cancers including adenocarcinoma,
CC leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,
CC breast, cervix or prostate). Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3932 BP; 1089 A; 883 C; 775 G; 1185 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,71e-254 Length: 3932
Score: 259.00 Matches: 259
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 90.88% Indels: 0
DB: 7 Gaps: 0

US-09-889-733B-2 (1-285) x ABZ35939 (1-3932)

QY 27 GlnHisAlaLeuArgSerAlaLysGlyGluValAspValAlaArgAlaGluArgGln 46
Db 95 CAGCAGCCCTGAGAGAGCCGAGGCGGAGAGTGGAGTCCGCCGCGGACGCGAG 154
QY 47 HisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValGluLeuPro 66
Db 155 CACCAGCTCTACGTGGCGCTGTGGGCGACAGAGTGGGGCTGCAGGTGGTGGAGCTGCCG 214
QY 67 AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGlu 86
Db 215 GCGGACGAGAGCCCTCCGAGCTGCGTCTTCGAGAGAGCTGCGCGCTGTGCGAGGAG 274
QY 87 ThrAlaLeuIleThrArgProGlyAlaProSerArgLysGluValAspMetMetLys 106
Db 275 ACGGCCCTCATCACCGACCCCGGGCGCGAGCCGAGGAGGAGTTGACATGATGAA 334
QY 107 GluAlaLeuGluLysLeuGlnLeuAsnLeuValGluMetLysAspGluAsnAlaThrLeu 126
Db 335 GAAGCATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAGATGAAATCCACTTTA 394
QY 127 AspGlyGlyValLeuPheThrGlyArgGluPheValGlyLeuSerLysArgThr 146

Db 395 GATGGCGAGATGTTTATTACAGGCGAGAGAAATTTTGTGGGCTTTTCCAAAGGACA 454
 Qy 147 AsndlnargGlyValAGluLeuAlaLeuAlaAspThrPheLysAsnTyAlaValSerThrVal 166
 Db 455 AATCAACAGAGTCTGAAATCTTGCTGATGATCTTTTAAAGCACTATGACATCTCCACAGTG 514
 Qy 167 ProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuLeu 186
 Db 515 CCAAGTGGCAGATGGTTGCAATTTGAAGAGTTTCTGCAGCATGCTGGGCTTAACCTGATC 574
 Qy 187 AlaileGlySerSerGlySerAlaGluLeuAlaLeuLysValMetClnMetSerAsp 206
 Db 575 GCAATGGGCTAGTGAATCTGCAGAGAGCCCTTAAATCATGCAACAGATGATGATGAC 634
 Qy 207 HisArgTyrAspLysLeuThrValProAspAspPileAlaAlaAsnCysIleTyrLeuAsn 226
 Db 635 CACCGCTACGACAACTCACTGCTGCTGATGACATAGCAGCAAACTGATATATATCTAAAT 694
 Qy 227 IleProAsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAla 246
 Db 695 ATCCCCAACAAAGGCGACGCTTCTGCTGCCAGCCCGAGAGATATCCAGAAAGTGCA 754
 Qy 247 LysValTyrGluLysLeuLysAspHisMetLeuLeuProValSerMetSerGluLeuGlu 266
 Db 755 AAGGTTTATGAGAACTTGAGGACCATATGCTGATCCCGCTGAGCATGCTGAAGTGA 814
 Qy 267 LysValAspGlyLeuLeuThrCysCysSerValLeuLeuAsnLysLysValAspSer 285
 Db 815 AAGTGATGGTGGTCTCACCTGCTGCTGATTTTATTAACAGAAAGTAGACTCC 871
 RESULT 6
 AAL41494
 ID AAL41494 standard; DNA; 4208 BP.
 XX AAL41494;
 AC AAL41494;
 XX
 DT 21-NOV-2002 (first entry)
 XX
 DE Drug metabolising enzyme encoding DNA - 7671089CBI.
 KW Anti-HIV; antiarteriosclerotic; dermatological; cytostatic; thyromimetic;
 KW osteoporosis; thrombolytic; ophthalmologic; antilipaseic; hepatotropic;
 KW antidiarrhoeic; antiinflammatory; virucide; immunogenic; autoimmune;
 KW drug metabolising enzyme; DME; inflammatory; AIDS; atherosclerosis;
 KW contact dermatitis; cell-proliferative; cancer; cirrhosis; dwarfism;
 KW developmental; hypothyroidism; endocrine; osteoporosis; thrombosis;
 KW diabetes; glaucoma; keratitis; metabolic; hyperlipidaemia; diarrhoea;
 KW cystic fibrosis; gastrointestinal; gastroenteritis; liver; hepatitis;
 KW Reye's syndrome; exogenous compound; gene therapy; enzyme; human; gene;
 ds.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 200..1225
 FT /*tag= a
 FT /product= "Drug metabolising enzyme protein"
 XX
 PN W020026654-A2.
 XX
 PD 29-AUG-2002.
 XX
 PF 14-FEB-2002; 2002WO-US004918.
 XX
 XX 16-FEB-2001; 2001US-0269643P.
 PR 23-FEB-2001; 2001US-0271332P.
 PR 16-MAR-2001; 2001US-0276767P.
 PR 06-APR-2001; 2001US-0282077P.
 PR 19-APR-2001; 2001US-0285447P.
 PR 27-APR-2001; 2001US-0287060P.
 PR 03-MAY-2001; 2001US-0288543P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.

XX
 PI Astromoff A, Au-Young J, Baughn MR, Ding L, Duggan BM;
 PI Forsythe IU, Giesen KJ, Griffin JA, Lee EA, Lu Y, Richardson TW;
 PI Ring HZ, Sanjanwala MM, Swarnakar A, Wallia NK, Warren BA, Xu Y;
 PI Yue H, Zebajadian Y;
 XX
 DR WPI; 2002-674949/72.
 DR P-PSDB; AAO22798.
 XX
 PT New drug metabolizing enzymes (DME) useful for diagnosing, treating and
 PT preventing diseases or conditions associated with aberrant DME
 PT expression, e.g. cancer, AIDS, atherosclerosis, diabetes, glaucoma,
 PT hepatitis, osteoporosis.
 XX
 PS Claim 5; Page 164-165; 166pp; English.
 PS
 CC The invention relates to an isolated polypeptide comprising 12 fully
 CC defined sequences of 81-615 amino acids given in the specification; a
 CC naturally occurring amino acid sequence at least 50% identical to, having
 CC 81-599 amino acids, at least 93% identical to a sequence of 529 amino
 CC acids, or at least 97% identical to a sequence of 615 amino acids, all
 CC given in the specification; or a biologically active or immunogenic
 CC fragment of the polypeptide. The polypeptides and polynucleotides are
 CC useful in diagnosing, treating and preventing diseases or conditions
 CC associated with the decreased expression or overexpression of a drug
 CC metabolising enzyme (DME), such as autoimmune/inflammatory (e.g. AIDS,
 CC atherosclerosis, contact dermatitis) cell-proliferative (e.g. cancer,
 CC cirrhosis), developmental (e.g. dwarfism, hypothyroidism), endocrine
 CC (e.g. osteoporosis, thrombosis, diabetes), eye (e.g. glaucoma,
 CC keratitis), metabolic (e.g. hyperlipidaemia, cystic fibrosis),
 CC gastrointestinal (e.g. gastroenteritis, diarrhoea), or liver (e.g.
 CC hepatitis, Reye's syndrome) disorders. These are also useful in assessing
 CC the effects of exogenous compounds on the expression of nucleic acid and
 CC amino acid sequences of DME. The DME or its fragments are useful in
 CC screening compounds for effectiveness as agonist or antagonist of the
 CC polypeptides, or in altering the expression of the target polynucleotide
 CC and compounds that specifically bind to or modulate the activity of the
 CC polypeptide. The microarray is useful in monitoring or measuring protein-
 CC protein interactions, drug-target interactions, and gene expression
 CC profiles. This polynucleotide of the invention can be used in gene
 CC therapy. This polynucleotide sequence represents the DNA encoding a drug
 CC metabolising enzyme of the invention
 XX
 SQ Sequence 4208 BP; 1092 A; 1017 C; 883 G; 1216 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 6,09e-254 Length: 4208
 Score: 259.00 Matches: 259
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 90.88% Indels: 0
 DB: 6 Gaps: 0
 US-09-889-733B-2 (1-285) x AAL41494 (1-4208)
 Qy 27 GlnHisAlaLeuArgSerAlaLysGlyGluValAlaAspValAlaArgAlaGln 46
 Db 446 CAGCAGCGCTCAGAGCGCCAGGCGAGGAGTGGACGTCCGCCGCGGACGCGAG 505
 Qy 47 HisGlnLeuTyrValGlyValLeuLysSerLysLeuGlnValValGluLeuPro 66
 Db 506 CACCAGCTCTACGTGGGCGTCTGGGCGAGCAAGTGGCGCTGCGAGCTGCCG 565
 Qy 67 AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGlu 86
 Db 566 GCCGACGAGAGCTTCCGACTGCGTCTTCGTGGAGCAGCTGGCGGTGTGCGAGGAG 625
 Qy 87 ThrAlaLeuIleThrArgProGlyAlaProSerArgLysGlyValAlaAspMetLys 106
 Db 626 ACGGCCCTCATCACCGACCCCGGGCGCCGAGGAGGAGGTGATGATGATGAAA 685
 Qy 107 GluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAlaThrLeu 126

Db 686 GAAGCATTAAGAAACTTCACTAGCTCAATATAGTAGATGAAGATGAAGAAATGCAACTTTA 745
 Qy 127 AspGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLysArgThr 146
 Db 746 GATGGCGGAGATGTTTATTACAGCAGCAGAGATTTTGTGGGCTTTCCAAAGGACCA 805
 Qy 147 AsnGlnArgGlyAlaGluLeuLeuAlaAspThrPheLysAspTyrAlaValSerThrVal 166
 Db 806 AATCAACAGAGTCTGTAATCTTGGCTGATACCTTTAAGGACTATGCACTCTCCACAGTG 865
 Qy 167 ProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProLeuLeuLeu 186
 Db 866 CCAGTGGCAGATGGTGTGCAATTTGAAGAGTTTCTGCAGCATGCTGGGCTAACTGATC 925
 Qy 187 AlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAsp 206
 Db 926 GCAATGGGTCTAGTGATCTGCACAGAGCGCTTTRAGATCATGCAACAGATGAGTGAC 985
 Qy 207 HisArgTyrAspLysLeuThrValProAspAspIleAlaAlaLeuLysLysLeuLeu 226
 Db 986 CACCGCTACGACAACTCACTGTGCTGATGACATAGCAGCAAACTGTATATATCTAAAT 1045
 Qy 227 IleProAsnLysGlyHisValLeuLeuHisArgThrProGluLysTyrProGluSerAla 246
 Db 1046 ATCCCCAACAAAGGCGACGCTTGTCTGCACCGAACCCCGAAGAGATCCAGAAAGTGCA 1105
 Qy 247 LysValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGlu 266
 Db 1106 AAGTTTATGAGAACTGAGGACCATATGCTGATCCCCGTGAGCATGCTGNACTGGAA 1165
 Qy 267 LysValAspGlyLeuLeuThrCysCysSerValLeuLeuAsnLysLysValAspSer 285
 Db 1166 AAGTGGATGGGTCTCACTGCTGCTCAGTTTAAATTAACAGAAAGTAGACTCC 1222
 RESULT 7
 ADE53981 standard; cDNA; 3778 BP.
 ID ADE53981
 AC ADE53981;
 XX
 XX 29-JAN-2004 (first entry)
 DE Human prostate cancer cDNA #328.
 XX
 KW Human; prostate cancer; ss; cDNA combination; differential expression;
 KW gene.
 OS Homo sapiens.
 XX
 PN US2003190640-A1.
 XX
 PD 09-OCT-2003.
 XX
 PF 29-MAY-2002; 2002US-00252157.
 XX
 XX 31-MAY-2001; 2001US-0295048P.
 PR
 XX (FARI/) FARIS M.
 PA (PEAR/) PEARSON C I.
 XX
 XX Faris M, Pearson CI;
 XX WPI; 2003-831619/77.
 XX
 PT New combination comprising cDNAs that are differentially expressed in
 PT prostate cancer, useful for diagnosing, treating or monitoring the
 PT progression of treatment of prostate cancer.
 XX
 XX Claim 1; SEQ ID NO 328; 42pp; English.
 XX
 XX The invention relates to a combination comprising a number of cDNAs
 CC expressed in prostate cancer. The invention also relates to a method for
 CC detecting differential expression of one or more cDNAs in a sample

CC containing nucleic acids by hybridising a substrate with the nucleic
 CC acids, thus forming one or more hybridisation complexes, detecting
 CC hybridisation complex formation and comparing the complexes formed with
 CC standard complexes, where differences between the standard and the sample
 CC complex formation indicate differential expression of cDNAs in the
 CC sample. The differential expression is diagnostic of prostate cancer. The
 CC invention also relates to proteins and antibodies related to the cDNAs.
 CC The combination is useful for diagnosing, treating or monitoring the
 CC progression of treatment of prostate cancer. The antibodies are useful
 CC for detecting prostate cancer. This sequence represents a human prostate
 CC cancer cDNA of the invention.
 XX

SQ Sequence 3778 BP; 1064 A; 821 C; 719 G; 1150 T; 0 U; 24 Other;

Alignment Scores:

Pred. No.: 2,89e-219 Length: 3778
 Score: 225.00 Matches: 238
 Percent Similarity: 59.58% Conservative: 0
 Best Local Similarity: 59.58% Mismatches: 0
 Query Match: 78.95% Indels: 1
 DB: 9 Gaps: 0

US-09-889-733B-2 (1-285) x ADE53981 (1-3778)

Qy	48	GlnLeuTyrValGlyValLeuGlySerLysLeuGlnValValGluLeuProAla	67
Db	2	CAGCTCTACGTGGGCGTCTGGCGAGCAAGCTGGGGCTGCAGGTGGTGGAGCTGCCGCC	61
Qy	68	AspGluSerLeuProAspCysValPheValGluAspValAlaValCysGluGluThr	87
Db	62	GACGAGAGCCCTCCGGACTGCGTCTTCTGTGGAGGACGTGGCGGTGTGGGAGGAGCG	121
Qy	88	AlaLeuIleThrArgProGlyValAlaProSerArgArgLysGluValAspMetLysGlu	107
Db	122	GCCTCATCACCGACCCGCGCGCGCGAGCGAGGAGGAGGTTGACATGATGAAGAA	181
Qy	108	AlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeuAsp	127
Db	182	GCATTTAGAAAACTTTCAGCTCAATATAGTAGAGTGAAGATGAATGCAATTTAGAT	241
Qy	128	GlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLysArgThrAsn	147
Db	242	GGCGGAGATGTTTTATTTCACAGGCAGAGAAATTTTTTGTGGGCTTTCCAAAGGACAAAT	301
Qy	148	GlnArgGlyAlaGluLeuLeuAlaAspThrPheLysAspTyrAlaValSerThrValPr	167
Db	302	CCACACGAGTGTGAAATCTTGGCTGATACCTTTAAGGACTATGCGAGTCTCCACAGTGCC	361
Qy	167	OValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAl	187
Db	362	AGTGGCAGATGGGTGCAATTTGAAGAGTTTCTGCAGCATGCTGGGCTAACTGATCGC	421
Qy	187	AlleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHi	207
Db	422	AATGGGTCTAGTGAATCTGCACAGAGGCGCTTAAAGATCATGCAACAGATGAGTGACA	481
Qy	207	ArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIle	227
Db	482	CCGCTACGACAAAATCACTGTCCTGATGACATAGCAGCAAACTGTATATATCTAAATAT	541
Qy	227	eProAsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAlaLys	247
Db	542	CCCCAACAAAGGCGACGCTTGTGCTGCACCGAACCCCGAAGAGTATCCAGAAAGTGCAAA	601
Qy	247	sValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGluLys	267
Db	602	GTTTTATGAGAACTGAAGGACCATATGCTGATCCCCGTGAGCATGTCTGAACCTGAAAA	661
Qy	267	sValAspGlyLeuLeuThrCysCysSerValLeuLeuAsnLysLysValAspSer	285
Db	662	GGTGGATGGGTGCTCACCTGCTGCTCAGTTTAAATTAACAGAAAGTAGACTCC	716

RESULT 8

Db 684 GAGGAGACCGCCTCATCAACCCGCGCGCGCGCTAGCCGACGAGGAGGTTGACATG 743
QY MetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAla 124
Db 744 ATGAAAGAGCGCTTTGAAATACTTCAGCTCAACATAGTAGACATGAAAGATGAAATGCA 803
QY 125 ThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLys 144
Db 804 ACCTTAGATGGTGGGACGCTCTATTCACAGCAGAGAGTCTTTTGTGGGCTTTCCAAA 863
QY 145 ArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSer 164
Db 864 AGACAATCAACAGAGGTCTGAGATCTTGGCTGATCTTCAAGGACTACGACGTTTCC 923
QY 165 ThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAs 184
Db 924 ACAGTCCCGCTGGCGGATTC-TTTGCATTTAAAGAGTTTCTGCAGCATGGCTGGCCCAA 982
QY 184 LeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMe 204
Db 983 CTGTGCGCAATAGGTCAGTGAATCTGCGCAGAGGCGCTCAAGATCATGCAACAGAT 1042
QY 204 tSerAspHisArgTyrAspLysLeuThrValProAspAsp 217
Db 1043 GAGTGACCCCGTTATGACAACTCACTGTACCGGACGAC 1082
RESULT 12
ABQ88101/c
ID ABQ88101 standard; cDNA; 83946 BP.
XX AC ABQ88101;
XX DT 18-SEP-2002 (first entry)
XX DE Human osteoblast differentiation related cDNA SEQ ID NO 8.
XX KW Human; osteoblast; stem cell differentiation; bone tissue deposition;
XX OS Homo sapiens.
XX XX WO200250301-A2.
XX PD 27-JUN-2002.
XX PF 18-DEC-2001; 2001WO-US048276.
XX PR 18-DEC-2000; 2000US-0255882P.
XX PR 24-APR-2001; 2001US-0285691P.
XX XX (GENE-) GENE LOGIC INC.
XX PA (PROC) PROCTER & GAMBLE CO.
XX PI Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
PI Mertz L;
XX XX WPI; 2002-557663/59.
XX XX Use of genes and their expression profiles associated with osteoblast
PT differentiation for screening modulators bone formation, for diagnosing
PT or treating e.g. osteoporosis, or as markers for the differentiation
PT process.
XX XX Claim 1; SEQ ID NO 8; 78pp + Sequence Listing; English.
XX XX The invention relates to genes and their expression profiles are used
CC for: (a) screening modulators of precursor stem cell differentiation into
CC osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
CC deposition of bone tissue, abnormal rate of osteoblast formation or
CC osteoporosis; or (c) treating or monitoring treatment of the conditions
CC cited in (b); or monitoring the progression of bone tissue deposition.
CC Specific conditions include postmenopausal osteoporosis, glucocorticoid
CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-

CC induced abnormalities in bone formation or bone loss, conditions that
CC involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
CC skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
CC or fibrous dysplasia. The present sequence is that of an osteoblast
CC differentiation associated cDNA marker of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 83946 BP; 26139 A; 16620 C; 16296 G; 24891 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3.08e-38 Length: 83946
Score: 49.00 Matches: 49
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 17.19% Indels: 0
DB: Gaps: 0
US-09-889-733B-2 (1-285) x ABQ88101 (1-83946)
QY 200 IleMetGlnMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAla 219
Db 18610 ATCATGCAACAGATGAGTGACCCGCTACGACAACTCACTGTGCTGATGACATAGCA 18551
QY 220 AlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrPro 239
Db 18550 GCAAACTGTATATATCTAAATATCCCAACAAAGGCGACCTTGTGTCACCGAACCCCG 18491
QY 240 GluGluTyrProGluSerAlaLysVal 248
Db 18490 GAGAGTATCCAGAAAGTGCAAGGTA 18464
RESULT 13
ABAI4642
ID ABAI4642 standard; DNA; 6373 BP.
XX AC ABAI4642;
XX XX 23-JAN-2002 (first entry)
XX DT Human nervous system related polynucleotide SEQ ID NO 6973.
XX DE Human; neurotropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antitickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebrotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX OS Homo sapiens.
XX XX WO200159063-A2.
XX PN 16-AUG-2001.
XX PD 17-JAN-2001; 2001WO-US001334.
XX PF 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205151P.
XX PR 07-JUN-2000; 2000US-0209467P.
XX PR 28-JUN-2000; 2000US-0214886P.
XX PR 30-JUN-2000; 2000US-0215135P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-0216880P.
XX PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226275P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.

PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-541565/60.

Nucleic acids encoding 3224 human nervous system antigen polypeptides,
useful for preventing, diagnosing and/or treating nervous system cancers
and metastases.

Disclosure; SEQ ID NO 6973; 1701pp + Sequence Listing; English.

The invention relates to novel genes (ABAI004-ABA21534) and proteins
(ABBI4678-ABBI8001) useful for preventing, treating or ameliorating
medical conditions e.g. by protein or gene therapy. The genes are
isolated from a range of human tissues disclosed in the specification.
The nucleic acids, proteins, antibodies and (ant)agonists are useful in
the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
disorders e.g. Addison's disease, allergies, autoimmune haemolytic
anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)

CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 6373 BP; 1810 A; 1396 C; 1180 G; 1987 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4,42e-28 Length: 6373
 Score: 38.00 Matches: 38
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 13.33% Indels: 0
 DB: 5 Gaps: 0

US-09-889-733B-2 (1-285) x ABA14642 (1-6373)

QY 211 LysLeuThrValProAspAspIleAlaAlaAsnCyseIleTyLeuAsnIleProAsnLys 230
 DB 2 AACTCACTGCTGCTGATGACATAGCAGCAAACTGTATATCTAATATATCCCAACAA 61
 QY 231 GlyHisValLeuLeuHisArgThrProGluGluTyProGluSerAlaLysVal 248
 DB 62 GGGCACGCTCTTGTGCACCGAACCAGGAGATATCCGAAAGTGCAGAGGTA 115

RESULT 14

ABA14641

ID ABA14641 standard; DNA; 6381 BP.

AC ABA14641;

XX 23-JAN-2002 (first entry)

XX Human nervous system related polynucleotide SEQ ID NO 6972.

XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
 KW antiparkinsonian; antischlicking; antianemic; antiarthritic; cancer;
 KW antirheumatic; hepatotropic; cerebrotective; antiinflammatory;
 KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
 KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

XX Homo sapiens.

XX WO200159063-A2.

XX 16-AUG-2001.

XX 17-JAN-2001; 2001WO-US001334.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184644P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209457P.

XX 28-JUN-2000; 2000US-0214866P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217497P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235835P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 20-OCT-2000; 2000US-0242221P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246529P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249267P.
PR 17-NOV-2000; 2000US-0249268P.
PR 17-NOV-2000; 2000US-0249269P.
PR 17-NOV-2000; 2000US-0249270P.
PR 17-NOV-2000; 2000US-0249271P.
PR 17-NOV-2000; 2000US-0249272P.
PR 17-NOV-2000; 2000US-0249273P.
PR 17-NOV-2000; 2000US-0249274P.
PR 17-NOV-2000; 2000US-0249275P.
PR 17-NOV-2000; 2000US-0249276P.
PR 17-NOV-2000; 2000US-0249277P.
PR 17-NOV-2000; 2000US-0249278P.
PR 17-NOV-2000; 2000US-0249279P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 06-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254037P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-541565/60.

Nucleic acids encoding 3224 human nervous system antigen polypeptides, useful for preventing, diagnosing and/or treating nervous system cancers and metastases.

Disclosure; SEQ ID NO 6972; 1701pp + Sequence Listing; English.

The invention relates to novel genes (ABAI1004-ABA21534) and proteins (ABE14678-ABE18001) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer, and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological disorders such as cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 6381 BP; 1809 A; 1396 C; 1182 G; 1994 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4-43e-28 Length: 6381
Score: 38.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.33% Indels: 0
DB: 5 Gaps: 0

US-09-889-733B-2 (1-285) x ABA14641 (1-6381)

QY 211 LysLeuThrValProAspAspIleAlaAlaGlyCysIleTyrLeuAsnIleProHisLys 230
DB 2 AAATCACTGTCCTGATGACATAGCAGCAAACTATATATCTAAATATCCCAACAAA 61

QY 231 GlyHisValLeuHisArgThrProGluGluTyrProGluSerAlaLysVal 248

DB 62 GGGCAGCTTGTCTGCACCGAACCCCGGAGAGTATCCAGAAAGTGCAAGGTA 115

RESULT 15

AAC10265

ID AAC10265 standard; cDNA; 420 BP.

XX

AC AAC10265;

XX

DT 06-OCT-2000 (first entry)

XX

DE Human secreted protein 5' EST, SEQ ID NO: 14340.

XX

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

XX

OS Homo sapiens.

XX

PN EP1033401-A2.

XX

PD 06-SEP-2000.

XX

PF 21-FEB-2000; 2000EP-00200610.

XX

PR 26-FEB-1999; 99US-0122487P.

XX

PA (GEST) GENSET.

XX

PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX

WPI; 2000-500381/45.

XX

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX

PS Claim 1; SEQ ID NO 14340; 71pp + Sequence Listing; English.

XX

CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No CRF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors

XX

SQ Sequence 420 BP; 131 A; 95 C; 94 G; 97 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.: 5.54e-18 Length: 420
Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.47% Indels: 0
DB: 3 Gaps: 0

US-09-889-733B-2 (1-285) x AAC10265 (1-420)

QY	102	ValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAsp	121
DB	338	GTGACATGATGAAAGAGCATTAGAAAACTTCAGCTCAATATAGTAGAGATGAAGAT	397
QY	122	GluAsnAlaThrLeuAspGly	128
DB	398	GAAATGCACTTAGATGGC	418

Search completed: June 8, 2004, 01:20:35
Job time : 411 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

QM protein - nucleic search, using frame_plus_p2n model

Run on: June 8, 2004, 01:05:17 ; Search time 2554 Seconds
(without alignments)
3332.311 Million cell updates/sec

Title: US-09-889-733B-2
Perfect score: 285
Sequence: 1 MAGLGHPSAFGRATHAVRA.....EXVDGLTCCSVLNNKVD 285

Scoring table:
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues
Word size: 1

Total number of hits satisfying chosen parameters: 55025477

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DBV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09889733/runat_07062004_083557_8367/app_query.fasta_1.455
-DB=EST -QPM=fastap -SUFFIX=olip2n.rst -MINMATCH=0.1 -COPEL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFWT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09889733 @CGN 1.1 5180 @runat_07062004_083557_8367 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:
2: em_esthum:
3: em_estin:
4: em_estnu:
5: em_estov:
6: em_estpl:
7: em_estro:
8: em_hic:
9: gb_est1:
10: gb_est2:
11: gb_hic:
12: gb_est3:
13: gb_est4:
14: gb_est5:
15: em_estfun:
16: em_estom:
17: em_gss_hum:
18: em_gss_inv:
19: em_gss_pin:
20: em_gss_vrt:
21: em_gss_fun:
22: em_gss_mam:
23: em_gss_mus:
24: em_gss_pro:
25: em_gss_rod:
26: em_gss_pig:
27: em_gss_vrl:
28: gb_gss1:

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	171	60.0	522	9	AI751463	AI751463 cn10b04.x
2	151	53.0	889	12	BI763390	BI763390 603049879
3	150	52.6	455	29	AY406329	AY406329 Homo sapi
4	116	40.7	356	12	BG977444	BG977444 RCS-C1016
5	111	38.9	624	10	AW514143	AW514143 hq2a02.x
6	106	37.2	535	9	AU279491	AU279491 AU279491
7	106	37.2	656	13	BY742062	BY742062 BY742062
8	106	37.2	746	14	CB249709	CB249709 UI-M-EXO-
9	106	37.2	862	13	BQ899468	BQ899468 AGENCOURT
10	106	37.2	1001	13	BY709864	BY709864 BY709864
11	106	37.2	1109	11	AK010430	AK010430 Mus muscu
12	102	35.8	760	14	CA512485	CA512485 UI-R-RJO-
13	93	32.6	702	13	BY731938	BY731938 BY731938
14	87	30.5	308	10	BF818963	BF818963 RCS-C1016
15	82	28.8	455	29	AY406330	AY406330 Pan trogl
16	80	28.1	685	14	CB247154	CB247154 UI-M-EIO-
17	78	27.4	777	14	CD349526	CD349526 UI-M-EYO-
18	77	27.0	385	14	CB702741	CB702741 AMGNNUC:S
19	76	26.7	556	12	BM964027	BM964027 UI-M-RQO-
20	75	26.3	406	9	AI017468	AI017468 uc3dc07.x
21	73	25.6	555	9	AI050531	AI050531 uc3dc07.y
22	72	25.3	731	10	BB611533	BB611533 BB611533
23	70	24.6	501	29	CG617513	CG617513 OST311187
24	69	24.2	776	14	CB317151	CB317151 AGENCOURT
25	68	23.9	806	9	AU051351	AU051351 AU051351
26	67	23.5	276	10	BF818967	BF818967 RCS-C1016
27	66	23.2	507	14	AI225556	AI225556 uj05e03.y
28	65	22.8	438	14	CB751364	CB751364 AMGNNUC:N
29	63	22.1	498	9	AI930353	AI930353 uc3f11.y
30	61	21.4	550	9	AA986603	AA986603 uc3f11.y
31	52	18.2	645	29	CG477727	CG477727 OST7734.M
32	52	18.2	733	14	CA750500	CA750500 UI-M-EYO-
33	51	17.9	432	10	BE864630	BE864630 UI-M-BH1-
34	50	17.5	450	14	CB787749	CB787749 AMGNNUC:N
35	50	17.5	654	9	AU296050	AU296050 AU296050
36	49	17.2	338	13	BY331431	BY331431 BY331431
37	48	16.8	639	13	BQ399610	BQ399610 NISC.mp04
38	47	16.5	486	9	AI463821	AI463821 va31a06.y
39	46	16.1	343	9	AA403666	AA403666 va31a06.y
40	46	16.1	455	29	AY406331	AY406331 Mus muscu
41	46	16.1	530	9	AI930086	AI930086 uc3f11.x
42	46	16.1	547	9	AI661738	AI661738 va31a06.x
43	46	16.1	548	14	CD774733	CD774733 UI-M-AQO-
44	46	16.1	569	9	AI265716	AI265716 uj05e03.x
45	46	16.1	571	14	CB522710	CB522710 UI-M-GKO-

ALIGNMENTS

RESULT 1
AI751463
LOCUS cn10b04.x1 Normal Human Trabecular Bone Cells Homo sapiens cDNA
DEFINITION clone NHTBC-cn10b04 random, mRNA sequence.
ACCESSION AI751463
VERSION AI751463.1 GI:5129727
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (Bases 1 to 522)


```

Db      300  GCCGACGAGAGCGCTCCCGACTCGCTCTTCGTGGAGGAGCGTGGCGGTGGTGGAGGAG 359
Qy      87  ThrAlaLeuIleThrArgProGlyAlaProSerArgLysGluValAspMetMetLys 106
Db      360  ACGGCCCTCATCCCGACCCGGCGCGCGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 419
Qy      107  GluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeu 126
Db      420  GAAGCATTTAGAAAACCTTCAGCTCAATATAGTAGAGATGAAGATGAAGATGAAGATGAAG 479
Qy      127  AspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLysArgThr 146
Db      480  GATGGCGGAGAGTGTATATCACAGCAGAGAAATTTTGTGGGCTTTCCTCAAGAGGACA 539
Qy      147  AsnGlnArgGlyAlaGluLeuAlaAspThrPheLysAspTyAlaValSerThrVal 166
Db      540  AATCAACGAGGTGCTGAAATCTTGGCTGATCTTTAAGGACTATGCACTCTCCACAGTG 599
Qy      167  ProValAlaAspGlyLeuHisLeuLysSerPhe 177
Db      600  CCAGTGGCAGATGGGTTCATTGAAGAGTTTT 632

RESULT 3
AY406329
LOCUS      455 bp      DNA      linear      GSS 15-DEC-2003
DEFINITION Homo sapiens DDH1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION  AY406329
VERSION     AY406329.1  GI:39762303
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 455)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE       Inferring nonneutral evolution from human-chimp-mouse orthologous
genes
JOURNAL     Science 302 (5652), 1960-1963 (2003)
PUBMED     14671302
REFERENCE   2 (bases 1 to 455)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE       Direct Submission
JOURNAL     Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES    Location/Qualifiers
source      1..455
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            <1..>455
            /gene="DDH1"
            /locus_tag="HCM2521"

gene
Alignment Scores:
Pred. No.:      6.2e-136      Length:      455
Score:          150.00      Matches:      150
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    52.63%      Indels:      0
DB:             29      Gaps:      0

ORIGIN
US-09-889-733b-2 (1-285) x AY406329 (1-455)

```

```

Qy      136  ArgGluPheValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuAla 155
Db      3  AGAGAATTTTGTGGGCTTTCCAAAAGGACAAATCAACGAGGTGCTGAAATCTTGGCT 62
Qy      156  AspThrPheLysAspTyAlaValSerThrValProValAlaAspGlyLeuHisLeuLys 175
Db      63  GATACCTTTTAAAGACATATGCAGTCTCCACAGTGCCAGTGCCAGATGGTTGCAATTTGAAG 122
Qy      176  SerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGln 195
Db      123  AGTTTCTGCAGCATGGCTGGGCTTAACCTGATCGCAATTCGGTCTAGTGAATCTGCACAG 182
Qy      196  LysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyTrAspLysLeuThrValPro 215
Db      183  AAGCCCTTTAGATCATGCAACAGATGATGACCCGCTACGACAACTCAGTGTGCT 242
Qy      216  AspAspIleAlaAsnCysIleTyTrLeuAsnIleProAsnLysGlyHisValLeuLeu 235
Db      243  GATGACATAGCAGCAAACTGTATATATCTAAATATATCCCAACAAAGGCGACGCTTCTGCTG 302
Qy      236  HisArgThrProGluGluTyProGluSerAlaLysValTyTrGluLysLeuLysAspHis 255
Db      303  CACGAAACCCCGAAGAGTATCCAGAAAGTGCAAAAGGTTTATGAGAACTGAGGACCAT 362
Qy      256  MetLeuIleProValSerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCys 275
Db      363  ATGCTGATCCCGTGAGCATGTCTGAACTGGAAAAGTGATGGGCTGCTCACCTGTGCTG 422
Qy      276  SerValLeuLeuLeuLysLysValAspSer 285
Db      423  TCAGTTTAAATTAACAAGAAAGTAGACTCC 452

RESULT 4
BG977444
LOCUS      356 bp      mRNA      linear      EST 12-JUN-2001
DEFINITION RCS-CT0161-040101-013-D07 C10161 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BG977444
VERSION     BG977444.1  GI:14380179
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 356)
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE       Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED     20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the PAPSP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?l=ACSt2-RC5-CI0161-
            040101-013-D07&t3=2001-01-04&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 356.
            Location/Qualifiers
            1..356
            /organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="C10161"
/note="Organ: colon ins; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORSTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Alignment Scores:
Pred. No.: 7,088-103 Length: 356
Score: 116.00 Matches: 116
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 40.70% Indels: 0
DB: 12 Gaps: 0

US-09-889-733B-2 (1-285) x BG977444 (1-356)

QY 37 GluValAspValAlaArgAlaGluArgGlnHisGlnLeuTyValGlyValLeuGlySer 56
Db 7 GAGGTGGACGTCGCCCGCGGAGCGGAGCAGCAGCAGCTCTACGTGGCGGTCTGGCGAGC 66
QY 57 LysLeuGlyLeuGlnValValGluLeuProAlaAspGluSerLeuProAspCysValPhe 76
Db 67 AAGCTGGGGCTCGAGGTGGTGGAGCTGCCGCCGCGAGAGCCCTCCGACTGGCTCTTC 126
QY 77 ValGluAspValAlaValValCysGluGluThrAlaLeuLeuThrArgProGlyAlaPro 96
Db 127 GTGGAGGAGCTGGCGTGGTGGTGGCGAGAGCAGCGCCCTCATCCCGACCGCGGCGCGC 186
QY 97 SerArgArgLysGluValAspMetMetLysGluAlaLeuGluLysLeuGluLeuAsnIle 116
Db 187 ACCCGGAGGAGCGAGGTGGACATGATGAAGAAGCATTAGAAAACCTTCAGCTCAATATA 246
QY 117 ValGluMetLysAspGluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArg 136
Db 247 GTAGAGTGAAGAGTGAATGCCACTTTAGTGGCGGAGATGTTTTATTACAGGCAGA 306
QY 137 GluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGlu 152
Db 307 GAATTTTGTGGGCTTTCCAAAAGGACAAATCAACGAGGTGCTGAA 354

RESULT 5
AW514143/c 624 bp mRNA linear EST 03-MAR-2000
LOCUS hd24a02.x1 Soares NFL T GBC S1 Homo sapiens cDNA clone
IMAGE:2910410 3, similar to TR:094760 O94760 NG,NG-DIMETHYLARGININE
DIMETHYLAMINOHYDROLASE 1, mRNA sequence.

ACCESSION AW514143 GI:7152225
VERSION AW514143.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 624)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 346.
Location/Qualifiers
1. .624

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="C10161"
/lab_host="DH10B"
/clone_lib="Soares NFL T GBC S1"
/note="Organ: pooled; Vector: p77T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI-CCAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN

Alignment Scores:
Pred. No.: 9,98e-98 Length: 624
Score: 111.00 Matches: 111
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 38.95% Indels: 0
DB: 10 Gaps: 0

US-09-889-733B-2 (1-285) x AW514143 (1-624)

QY 175 LysSerPheCysSerMetAlaGlyProAsnLeuLeuAlaIleGlySerSerGluSerAla 194
Db 477 AAGAGTTTCTGACGATGGCTGGGCTTAACCTGATCGCAATTGGGCTAGTGNATCTGCA 418
QY 195 GlnLysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyrAspLysLeuThrVal 214
Db 417 CAGAGGCCCTTAAGATCATGCAACAGATGAGTGACCCGCTACGACAACTCACTGTG 358
QY 215 ProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeu 234
Db 357 CCTGATGACATACAGCAAACTGTATATATCTAAATATCCCCAACAAAGGCGACGCTTG 298
QY 235 LeuHisArgThrProGluGluTyrProGluSerAlaLysValTyrGluLysLeuLysAsp 254
Db 297 CTGACCGAACCCTGGAGAGATATCCAGAAAGTGCAAGGTTTATGAGAACTGAAGGAC 238
QY 255 HisMetLeuLeuProValSerMetSerGluLeuGluLysValAspGlyLeuLeuThrCys 274
Db 237 CATATGCTGATCCCGCTGAGCATGTCTGAACTGGAAGAGTGATGGCTGCTCACCTGC 178
QY 275 CysSerValLeuLeuLeuLysLysValAspSer 285
Db 177 TGCTCAGTTTTTAATTACAGAAAGTAGACTCC 145

RESULT 6
AU279491 535 bp mRNA linear EST 31-JUL-2003
LOCUS AU279491 CHONS1 Homo sapiens cDNA clone CHONS1000018', mRNA
DEFINITION AU279491 sequence.
ACCESSION AU279491
VERSION AU279491.1 GI:28298718
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 535)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Inabayashi, H., Mori, T., Gojo, S., Kiyono, T., Sugiyama, T., Irie, R.,
Iscgal, T., Kata, J., Tomoya, Y. and Umezawa, A.
Redifferentiation of dedifferentiated chondrocytes and
chondrogenesis of human bone marrow stromal cells via chondrosphere
formation with expression profiling by large-scale cDNA analysis
Exp. Cell Res. 286 (1), 35-50 (2003)

MEDLINE 22760698
 PUBMED 12878157
 COMMENT Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomics@hri.co.jp
 HRI human cDNA project, Sugiyama, T.; Wakamatsu, A.; Irie, R.;
 Umezawa, A.; Fukuma, M.; Kusakari, S.; Hata, J.; Ishii, S.; Yamamoto, J.;
 Isono, Y.; Saito, K.; Nakamura, Y.; Masuho, Y.; Nagai, K.; Isogai, T.
 HRI human cDNA project: cDNA library construction & 5'-end one
 pass sequencing: Helix Research Institute.

FEATURES

source Location/Qualifiers
 1..535
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CHONS1000018"
 /cell_type="chondrocytes"
 /clone_lib="CHONS1"
 /notes="Vector: pME18SFL3"

ORIGIN

Alignment Scores:
 Pred. No.: 6 37e-93 Length: 535
 Score: 106.00 Matches: 106
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 37.19% Indels: 0
 DB: 9 Gaps: 0

US-09-889-733B-2 (1-285) x AU279491 (1-535)

Qy 102 ValAspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAsp 121
 Db 218 GTTGACATGATGAAGAAGCATTAGAAAACCTTCAGCTCAATATAGATGAGATGAAGAT 277
 Qy 122 GluAsnAlaThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheValGly 141
 Db 278 GAAATCCACTTTAGATGGCGAGATGTTTATTCACAGCCAGAGAAATTTTGTGGGC 337
 Qy 142 LeuSerLysArgThrAsnGlnAArgGlyAlaGluLeuAlaAspThrPheLysAspTyr 161
 Db 338 CTUTCCAAAGGACAAATCAACAGGTGCTGAAATCTTGCTGATCTTTTAAGGACTAT 397
 Qy 162 AlaValSerThrValProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAla 181
 Db 398 GCAGTCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 457
 Qy 182 GlyProAsnLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMet 201
 Db 458 GGGCTTAACCTGATCGGATGGTCTAGTGAATCTGCACAGAGCCCTTAAGATCATG 517
 Qy 202 GlnGlnMetSerAspHis 207
 Db 518 CAACAGATGATGACCAC 535

RESULT 7
 BY742062
 LOCUS BY742062 656 bp mRNA linear EST 17-DEC-2002
 DEFINITION BY742062 RIKEN full-length enriched, 2 days neonate sympathetic ganglion Mus musculus cDNA clone 7120491n07 5', mRNA sequence.
 ACCESSION BY742062
 VERSION BY742062.1 GI:27167348
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 656)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojohori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Guatinchich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempie, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L. G., Wainwright, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 12354683
 12466851
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Saito-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp
 URL: http://genome.gsc.riken.go.jp/
 Fukuda, S., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Imotani, K., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Kondo, S., Konno, H., Koyama, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 RNA was provided by Dr. Akira Nakagawara (Div. of Biochemistry, Chiba Cancer Center Research Institute, 666-2 Nitona, Chuo-ku, Chiba, 260-8717 Japan) whose assistance is gratefully acknowledged. Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

FEATURES

source Location/Qualifiers
 1..656
 /organism="Mus musculus"
 /mol_type="mRNA"

```

/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="7120491N07"
/sex="mixed"
/tissue_type="sympathetic ganglion"
/dev_stage="2 days neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 2 days neonate
sympathetic ganglion"
/notes="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGATCCCAAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGATCCCAAGAGCTCTTTTCTTTTCTTTT 3']. cDNA
was cleaved with XhoI and BamHI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I.
Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.
-rRNA was provided by Akira Nakagawara, Div. of
Biochemistry, Chiba Cancer Center Research Institute,
666-2 Nitona, Chiba-ku, Chiba, 260-8717 Japan, whose
assistance we gratefully acknowledge."

ORIGIN
Alignment Scores:
Pred. No.: 7,986-93 Length: 656
Score: 106.00 Matches: 152
Percent Similarity: 98.70% Conservative: 0
Best Local Similarity: 98.70% Mismatches: 1
Query Match: 37.19% Indels: 2
DB: 13 Gaps: 0

US-09-889-733B-2 (1-285) x BY742062 (1-656)

QY 65 LeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAlaValCys 84
DB 12 CTGCCCGCGGACGAGAGCTCCCGACCTGCTGTTCTGTGGAGACGTCGCGCTGTCG 71
QY 85 GluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgGlyGluValAspMet 104
DB 72 GAGGAGAGCGGCCCTCATCACCGCCCGCGCGCCAGCCGAGAGAGGTTGACATG 131
QY 105 MetLysGluAlaLeuGluLysLeuGlnLeuSnlleValGluMetLysAspGluAsnAla 124
DB 132 ATGAAGAGGCTTTGGAAAAAATCTCAGCTCAACATAGTAGAGATGAAGATGAATGCA 191
QY 125 ThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLys 144
DB 192 ACTTTGGATGTTGGGAGCTCTTATTCACAGCGACAGAAATTTTGTGGCCITTCACAA 251
QY 145 ArgThrAsnGlnArgGlyAlaGluLeuAlaAspThrPheLysAspThrAlaValSer 164
DB 252 AGAACAAATCAACAGAGTGCTGAAATCTTGCTGATCTTTTAAAGGACTACGCGAGTCTCT 311
QY 165 ThrValProValAlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProAs 184
DB 312 ACAGTCCCTGTGGCGGATTC-TTTGCAATTAAGAGTTTCTGACGATGCGCGACCCAA 370
QY 184 LeuLeuAlaLeuGlySerSerGluSerAlaGlnLysLeuLysLeuMetGlnGlnMe 204
DB 371 CCTGATTGCAATAGGTCACAGCAATCTGCACAGAGCCCTCAAGATCATGCAACAGAT 430
QY 204 tSerAspHisArgTyrAspLysLeuThrValProAspAsp 217
DB 431 GAGTGACCATCGTTATGACAGCTCATCTGTTACCCGACGAC 470

CB249709 746 bp mRNA linear EST 15-JUL-2003
UI-M-EX0-by1-1-07-0-UI.r1 NIH_BMAP_EX0 Mus musculus cDNA clone
IMAGE:5719830 5', mRNA sequence.
CB249709
CB249709.1 GI:28389411
EST.
Mus musculus (house mouse)
Mus musculus
ORGANISM
REFERENCE
1 (bases 1 to 746)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
AUTHORS
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..746
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5719830"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_EX0"
/notes="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA
tail, is GTGCGTGGAA. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN
Alignment Scores:
Pred. No.: 9,218-93 Length: 746
Score: 106.00 Matches: 152
Percent Similarity: 98.70% Conservative: 0
Best Local Similarity: 98.70% Mismatches: 1
Query Match: 37.19% Indels: 2
DB: 14 Gaps: 0

US-09-889-733B-2 (1-285) x CB249709 (1-746)

QY 65 LeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAlaValCys 84
DB 30 CTGCCCGCGGACGAGAGCTCCCGACCTGCTGTTCTGTGGAGACGTCGCGCTGTCG 89
QY 85 GluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgArgGlyGluValAspMet 104
DB 90 GAGGAGACGCGCCCTCATCTACCCCGCGCGCGCCGAGGAGGAGGTTGACATG 149

```



```

QY 105 MetLysGluAlaLeuLysLeuGlnLeuAenIleValGluMetLysAspGluAenAla 124
DB 150 ATGAAGAAGAGCTTTGGAAAACCTTACGCTCAACATAGTAGATGAAGATGAATGCA 209
QY 125 ThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLys 144
DB 210 ACTTTGGATGGTGGGACGCTCTATTTCACAGCAGAGAAATTTTGTGGCCCTTCCAAA 269
QY 145 ArgThrAsnGlnArgGlyAlaGluLeuLeuAlaAspThrPheLysAspTyrAlaValSer 164
DB 270 AGAACAAATCAACAGAGGTCTGAATCTTGGGTGATCTATTTTAAAGACTACGACGCTCT 329
QY 165 ThrValProValAlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProAs 184
DB 330 ACAGTCCCTGTGGCGATTCTTTGCAATTAAGAGTTTCTCAGCATGGCCGGACCCAA 388
QY 184 nLeuLeAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMe 204
DB 389 CCTGATTGCAATAGGTCAGAGGCTCTGCACAGAAGGCCCTCAAGATCATGCAACAGAT 448
QY 204 tSerAspHisArgTyrAspLysLeuThrValProAspAsp 217
DB 449 GAGTGACCATCGTTATGACAAGCTCACTGTACCCGACGAC 488

RESULT 9
BQ899468
LOCUS BQ899468
DEFINITION AGENCOURT_8750357 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6332815
5', mRNA sequence.
ACCESSION BQ899468
VERSION BQ899468.1 GI:22911482
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 862)
NIH-MGC http://imgc.ncbi.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
Ph.D.
CDNA Library Preparation: ResGen, Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13791 row: a column: 08
High quality sequence stop: 704.
Location/Qualifiers
1. .862
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6332815"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 130"
/notes="Organ: oocytes; Vector: pCMV-SPORT6.1; Site 1:
EcoRV; Site 2: NotI; Cloned unidirectionally. Primer:
Oligo dt. Average insert size 1.95 kb. Constructed by
ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 1,08e-92 Length: 862
Score: 106.00 Matches: 152
Percent Similarity: 98.70% Conservative: 0
Best Local Similarity: 98.70% Mismatches: 1
Query Match: 37.19% Indels: 2

```

```

DB: 13 Gaps: 0
US-09-889-733B-2 (1-285) x BQ899468 (1-862)
QY 65 LeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAlaValCys 84
DB 323 CTGCCGCCAGAGAGCTGCCGACTGGCTGTTCGTGGAGGACGTGGCGCTCGTGTGC 382
QY 85 GluLysThrAlaLeuLysLeuThrArgProGlyValaProSerArgLysGluValaAspMet 104
DB 383 GAGGAGAGCGCCCTTCATCACCGCCCGGGCGCCCGCAGGAGGAGGTTGACATG 442
QY 105 MetLysGluAlaLeuGluLysLeuGlnLeuAenIleValGluMetLysAspGluAenAla 124
DB 443 ATGAAGAAGAGCTTTGGAAAACCTTACGCTCAACATAGTAGATGAAGATGAATGCA 502
QY 125 ThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLys 144
DB 503 ACTTTGGATGGTGGGACGCTCTATTTCACAGGACAGAAATTTTGTGGCCCTTCCAAA 562
QY 145 ArgThrAsnGlnArgGlyAlaGluLeuLeuAlaAspThrPheLysAspTyrAlaValSer 164
DB 563 AGAACAAATCAACAGAGGTCTGAATCTTGGGTGATCTATTTTAAAGACTACGACGCTCT 622
QY 165 ThrValProValAlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProAs 184
DB 623 ACAGTCCCTGTGGCGATTCTTTGCAATTAAGAGTTTCTCAGCATGGCCCGACCCAA 681
QY 184 nLeuLeAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMe 204
DB 682 CCTGATTGCAATAGGTCAGAGGCTCTGCACAGAAGGCCCTCAAGATCATGCAACAGAT 741
QY 204 tSerAspHisArgTyrAspLysLeuThrValProAspAsp 217
DB 742 GAGTGACCATCGTTATGACAAGCTCACTGTACCCGACGAC 781

RESULT 10
BQ709864
LOCUS BQ709864
DEFINITION BQ709864 RIKEN full-length enriched, ES cells Mus musculus cDNA
clone 2410006N07 5', mRNA sequence.
ACCESSION BQ709864
VERSION BQ709864.1 GI:27121074
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1001)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamakawa, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brusci, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.B., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Matais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Sempie, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlested, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,

```

Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL
Nature 420, 563-573 (2002)

MEDLINE
22354683

PUBMED
12466851

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

Location/Qualifiers
1..1001
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clones="2410006N07"
/cell_type="ES cells"
/lab_host="SOLR"

/clone_lib="RIKEN full-length enriched, ES cells"
/note="Site 1: XhoI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGATTCGAGTAAATAAATCCGCCCCCCCCC 3']". cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGATTCGAGTAAATAAATCCGCCCCCCCCC 3']".

ORIGIN

Alignment Scores:
Pred. No.: 1,286-92 Length: 1001
Score: 106.00 Matches: 152
Percent Similarity: 98.70% Conservative: 0

Best Local Similarity: 98.70% Mismatches: 1
Query Match: 37.19% Indels: 2
DB: 13 Gaps: 0

US-09-889-733B-2 (1-285) x BY709864 (1-1001)

QY 65 LeuProAlaAspGluSerLeuProAspCysValPheValGluAspValAlaValAlcys 84
DB 328 CTGCCCGCCGACGAGAGCTGCCGCTGTTCTGTTGGAGGACGTGGCGCTGTGTC 387
QY 85 GluGluThrAlaLeuLeuThrArgProGlyAlaProSerArgAlaGlyLysGluValAspMet 104
DB 388 GAGGAGAGCGCCCTCATCCCGCCCGGCGCCCGCCGAGGAGGAGTTGACATG 447
QY 105 MetLysGluAlaLeuGluLysGluLeuAsnLeuValGluMetLysAspGluAsnAla 124
DB 448 ATGAAGAGGCTTTGGAAAACTTCAGCTCAACATAGTAGAGATCAAGATGAATGCA 507
QY 125 ThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheValGlyLeuSerLys 144
DB 508 ACTTTGGATGTTGGGAGCTCTATTACAGCGAGAGATTTTTTTGGGGCTTTCCAA 567
QY 145 ArgThrAsnGlnArgGlyAlaGluLeuAlaAspThrPheLysAspTyrAlaValSer 164
DB 568 AGAACAAATCAACGAGGTGCTGAAATCTTGCTGATACTTTTAAGGACTAGCGACTCTCT 627
QY 165 ThrValProValAlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProAs 184
DB 628 ACAGTCCTCTGGGCGGATTC-TTTCATTTAAAGAGTTCTGCAGCATGGCCGACCCAA 686
QY 184 nLeuLeuAlaIleGlySerSerGlySerAlaGlnLysAlaLeuLysIleMetGlnMe 204
DB 687 CTGATTGCAATAGGTCCAGCGAATCTGCAGAGCGCTCAAGATCATGCAACAGAT 746
QY 204 tSerAspHisArgTyrAspLysLeuThrValProAspAsp 217
DB 747 GAGTGACCATCTGTTATGCAAGCTCACTGTATCCCGGAC 786

RESULT 11

AK010430

LOCUS

AK010430 1109 bp mRNA linear HTC 20-SEP-2003
Mus musculus ES cells cDNA, RIKEN full-length enriched library.
clone:2410006N07 product:dimethylarginine dimethylaminohydrolase 1,
full insert sequence.

ACCESSION

AK010430

VERSION

AK010430.1 GI:12845867

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

JOURNAL

99279253

MEDLINE

10349636

PUBMED

REFERENCE

2

AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL

20499374

MEDLINE

11042159

PUBMED

REFERENCE

3

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitaura, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsuoka, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Chara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system--394-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PubMed 11078861

REFERENCE
4. 15070551
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 585-690 (2001)
JOURNAL

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Group Phase I & II team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420. 563-573 (2002)

Nature 420, 563-573 (2002)

6 (bases 1 to 1109)

REFERENCE

AUTHORS

Adachi, J., Aizawa, K., Bono, H., Carninci, P., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Hara, A., Hayatsu, N., Hiramoto, K., Furuo, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tegawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, Y., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Sushiro-cho, Tsurumi-Ku, Yokohama
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGGATCCACAGAGCTCTTTTCTTTTCTT 3']. cDNA was prepared by using trisphalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGATCTCCGATTAATTAATTAATCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

FEATURES

```

1. .1109
/organism="Mus musculus"
/mol_type="mrna"
/strain="C57BL/6J"
/db_xref="FANTOM.DB:2410006N07"
/db_xref="MGI:1901307"
/db_xref="taxon:10090"
/clone="2410006N07"
/cell_type="ES cells"
/clone_lib="RIKEN full-length e
137. .994
CDS

```

CDs

```

137. .994
/note="unnamed protein product; dimethylarginine
dimethylaminohydrolase 1 (MGDI|GI:1916469, GB|AKO10430,
evidence: BLASTN, 100%, match=1109)
putative"
/codon_start=1
/protein_id="BAB26932.1"
/db_xref="GI:12845868"
/translation="MAGLHSPSAGRTATHVVRPPESLCRHLRRSQSEVDFARAE
RQHLYVGLGSLGVLQVLPADSPDCVFQEDVAVGCEETALTRPGAPSRKEV
DMWKEALEKLOINTVENKNDATLDGVDLFTGREFVGLSKETNORGAELIADTFKD

```

YAVSTVPVADSLHLKSFCSMAGPNLIAIGSSESAQKALKIMQMSDHRDYDKLTVPDDM
AANCIYLNIPSKGHVLLHRTPEEYEPESAKVEKLDHLLIPVNSEMEKVDGLLTCCS
VFINKIDS"

ORIGIN

Alignment Scores:	
Pred. No.:	1,43e-92
Score:	106.00
Percent Similarity:	98.70%
Best Local Similarity:	98.70%
Mismatches:	0
Query Match:	37.19%
Indels:	2
DB:	11
Gaps:	0
Length:	1,109
Matches:	152
Conservative:	1
Mismatches:	0
Indels:	2
Gaps:	0

US-09-889-733B-2 (1-285) x AK010430 (1-1109)

65 LeuProIlaAspGluSerLeuProAspCysValPheValGluAspValAlaValValCys 84

Dd	329	CTGCCCGCCGACGACGAGAGCTGCCCGACTGCCGTTTCGTGTGAGGACGCTGGCCGTCGTGTGTC	388
Qy	85	GluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMet	104

Db 389 GAGGAGACGGCCCTCATCACCCGCCCGGGGGCGCCAGCCGACGGAAGAGGTTGACATG 448

Qy
105 MetLysGluAlaIeuGluLysLeuGlnIeuAsnIleValIcIUMerLysAspGluAsnAla 124
|||
Db
449 ATGAACAGGGCTTTGGAAAAAACCITTCAGCTCAACATAGTAGAGATGAAGAATGAAAATGCA 508

Qy 125 ThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPheValGlyLeuSerLys 144

509	ACTTTGGATGTTGGGGACGCTCTATTTCACAGGCACAGAAATTTTTTGTGGGGCTTTTCCAAA	Db
145	ArgThrAsnGlnArgGluAlaGluIleLeuAlaAspThrPheIysAspTyrAlaValSer	Qy

569 AGAACAAATCAACGAGGTCGTAATCTTGGCTGATACATTTAAGGACTACGCAGTCTCT 628

Qy 165 ThrValProValAlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProAs 184
Db 629 ACAGTCCTCGTGCCCGATTC-TTTGCAITTAAGAGTTTCTGCAGCATGGCCGACCCAA 687

Qy 184 nLeu11leAla11leGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMe 204

Db 688 CCTGATTGCAATAGGGTCCAGCGAATCTGCACAGAAGGCCCTCAAGATCATGCAACAGAT 747

Qy 204 tSerAspHisArgTyrAspLysLeuThrValProAspAsp 217

D_b 748 GAGTGACCATCGTTATTGACAAGCTCAGTACCGACGAC 787

RESULT 12	760 bp	linear	EST 15-NOV-2000
CA512485		mRNA	
LOCUS	CA512485		

LOCUS	CA512485	760 bp	linear	03-12-2007
DEFINITION	UI-R-FJ0-cpz-1-06-0-UI.r1 UI-R-FJ0 Rattus norvegicus cDNA clone			
ACCESSION	CA512485	UI-R-FJ0-cpz-1-06-0-UI 5', mRNA sequence.		

ACCESSION
CAS12485
VERSION
CAS12485.1 GI:25003439
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)

SOURCE	ORGANISM
Rattus norvegicus (Norway rat)	Rattus norvegicus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE
1 (bases 1 to 760)
Rattus.
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
AUTHORS Ronald M. F. Leannon, G. and Soares, M. R.

AUTHORS	Bonaldo, M.P., Lennon, G. and Soares, M.B.
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL	Genome Res 6 (9) 791-805 (1996)

JOURNAL
Genome Res. 6 (9), 791-806 (1996)
MEDLINE
97044477
COMMENT
8889548
Contact: Sooyae MB

CONTACT: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
275 Neuroscience Road A156 MBDBB Iowa City IA 52242 USA

375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. James Lin, University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
 Seq primer: M13 REVERSE.

FEATURES

source
 Location/Qualifiers
 1..760
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-FJ0-cp2-1-06-0-UI"
 /tissue_type="embryo"
 /dev_stage="embryo"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-R-FJ0"
 /notes="vector: pfx-asc; Site 1: EcoR I; Site 2: Not I; UI-R-FJ0 is a cDNA library containing the following tissue(s): rat embryo. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CACTCTPACT. This library was created for the University of Iowa Program for Rat Gene Discovery and Mapping (Val Sheffield, Bento Soares and Tom Casavant)"

ORIGIN

Alignment Scores:
 Pred. No.: 7,52e-89 Length: 760
 Score: 102.00 Matches: 148
 Percent Similarity: 98.67% Conservative: 0
 Best Local Similarity: 98.67% Mismatches: 1
 Query Match: 35.79% Indels: 2
 DB: 14 Gaps: 0

US-09-889-733B-2 (1-295) x CA512485 (1-760)

Qy 69 GluSerLeuProAspCysValPheValGluAspValAlaValValCysGluGluThraAla 88
 Db 2 CAGAGCCTGCCTGACTGCGTTCGTGGAGACGTGCGCGTGTGCGGAGACCGCC 61
 Qy 89 LeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMetMetLysGluAla 108
 Db 62 CTCATCCCGCCCGCGCGCCCTTAGCCGACGAGAGGTTGACATGATGAAAGAGGCT 121
 Qy 109 LeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeuAspGly 128
 Db 122 TTGGAAAACCTCAGCTCAACATAGTAGAGATGAAGATGAAGATGAAGATGAAGATG 181
 Qy 129 GlyAspValLeuPheThrGlyArgGluPheValGlyLeuSerLysArgThrAsnGln 148
 Db 182 GGGGACGTCCTATTACAGGACGAGAGGTTTTTTGTGGCCCTTTCCAAAGACAAATCAA 241
 Qy 149 ArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSerThrValProVal 168
 Db 242 CGAGGTGCTGAGATCTTGGCTGATACTTTCAAGGACTACGAGTTTCCACAGTCCCGGTG 301
 Qy 169 AlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProLeuIleAla 188
 Db 302 GCGGATTC-TTGGCATTTAAAGAGTTTCTGAGATGCTGGCCCCAACCTGATCGCAAT 360
 Qy 188 eGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHisAr 208
 Db 361 AGGTCCTCAGTGATCTGGCGACAGGCGCTTCAGATCATGCAACAGATGAGTACCACCG 420

Qy 208 gTyxAspLysLeuThrValProAspAsp 217
 Db 421 TTATGACAAAGTCTACTGTACCGGACGAC 448

RESULT 13

BY731938

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BY731938 702 bp mRNA linear EST 17-DEC-2002
 BY731938 RIKEN full-length enriched, 8 cells embryo Mus musculus
 cDNA clone E860038F15 5', mRNA sequence.

BY731938 GI:27145065

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 702)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaudo, I., Otsu, N., Saito, R., Suzuki, H., Yamakawa, T.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Buit, C.,
 Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusci, V.,
 Chochova, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A.,
 Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,
 Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A.,
 Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
 Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,
 Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,
 Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K.,
 Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,
 Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
 Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
 Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukawa, S., Harai, A., Hashizume, W., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
 Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,
 Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

2234683

12466851

COMMENT

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Hirozane, T., Hori, F.,
 Fukuda, S., Hashizume, W., Hayashida, K., Kawai, J., Koizumi, Y.,
 Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Koizumi, Y.,
 Kondo, S., Konno, H., Koyama, S., Miyazaki, A., Murata, M., Nakamura, M.,
 Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, K., Sakazume, N.,
 Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,
 Takeda, Y., Waki, K., Watanabe, K., Watanabe, M. and Hayashizaki, Y.

Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source
 Location/Qualifiers
 1..702
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="B860038F15"
 /cell_type="8 cells"
 /dev_stage="8 cells embryo"
 /clone_lib="RIKEN full-length enriched, 8 cells embryo"

ORIGIN

Alignment Scores:
 Pred. No.: 4.17e-80 Length: 702
 Score: 93.00 Matches: 93
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 32.63% Indels: 0
 DB: 13 Gaps: 0

US-09-889-733B-2 (1-285) x BY731938 (1-702)

QY 65 LeuProAlaaspGluSerLeuProAspCysValPheValGluaspValAlaValValCys 84
 |||||
 Db 375 CTGCCCGCCGACGAGACCTGCGCGACTGCGTTCGTGGAGACGTCGCGCTGTGC 434
 |||||
 QY 85 GluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMet 104
 |||||
 Db 435 GAGGAGACGCGCCCTCATCCCGCCCGCCGCGCCGCGCCGCGCGAGGAGGTGACATG 494
 |||||
 QY 105 MetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAla 124
 |||||
 Db 495 ATGAAGAAGGCTTTGGAAACCTTCAGCTCACATAGTAGAAGATGAATGCA 554
 |||||
 QY 125 ThrLeuaspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLys 144
 |||||
 Db 555 ACTTTGGATGTTGGGACGCTCTATTACAGCGCAGAGATTTTGTGGCCCTTCCAAA 614
 |||||
 QY 145 ArgThrAsnGlnArgGlyAlaGluIleLeuAlaaspThr 157
 |||||
 Db 615 AGAACAAATCAACAGGTGCTGAAATCTTGCTGTACT 653
 |||||

RESULT 14

BP818963/c 308 bp mRNA linear EST 13-JAN-2001
 LOCUS R5-C10160-221200-012-A08 C10160 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BP818963
 ACCESSION BP818963.1 GI:12156591
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS 1. (bases 1 to 308)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.P., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed

TITLE

sequence tags.

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 PUBMED 10737800

COMMENT

Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC5&st2=RC5-CI0160-221200-012-A08&t3=2000-12-22&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 11
 High quality sequence stop: 308.

FEATURES

source
 Location/Qualifiers
 1..308
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="CI0160"
 /note="Organ: colon ins; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Alignment Scores:
 Pred. No.: 1.19e-74 Length: 308
 Score: 87.00 Matches: 87
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 30.53% Indels: 0
 DB: 10 Gaps: 0

US-09-889-733B-2 (1-285) x BP818963 (1-308)

QY 61 GlnValValGluLeuProAlaaspGluSerLeuProAspCysValPheValGluaspVal 80
 |||||
 Db 308 CAGGTGGTGGAGCTCCCGCCGCGCAGAGACCTTCGGACTCGCTTCGTGGAGACGCTG 249
 |||||
 QY 81 AlaValValCysGluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLys 100
 |||||
 Db 248 GCCGTGGTGTGGAGAGACGCGCCCTCATCCCGCCGCGCGCGCGCGGAGGAG 189
 |||||
 QY 101 GluValaspMetMetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLys 120
 |||||
 Db 188 GAGGTGACATGATGAAGAAGCATTAGAAAACCTTCAGCTCATATATAGTAGATGAAA 129
 |||||
 QY 121 AspGluAsnAlaThrLeuAspGlyGlyaspValLeuPheThrGlyArgGluPheVal 140
 |||||
 Db 128 GATGAATGCACTTTAGATGGCGGACATGTTTATTTCACAGCGCAGAGATTTTGTG 69
 |||||
 QY 141 GlyLeuSerLysArgThrAsn 147
 |||||
 Db 68 GGCCTTTCCAAAGGACAAAT 48
 |||||

RESULT 15

AY406330 455 bp DNA linear GSS 15-DEC-2003
 LOCUS Pan troglodytes DD4H1 gene, VIRTUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY406330
 VERSION AY406330.1 GI:39762304
 KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
AUTHORS 1 (bases 1 to 455)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D., and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source
1..455
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>455
/gene="DDAH1"
/locus_tag="HCM2521"

ORIGIN

Alignment Scores:
Pred. No.: 1.39e-69 Length: 455
Score: 82.00 Matches: 149
Percent Similarity: 98.68% Conservative: 0
Best Local Similarity: 98.68% Mismatches: 1
Query Match: 28.77% Indels: 2
DB: 29 Gaps: 0

US-09-889-733B-2 (1-285) x AY406330 (1-455)

QY 136 ArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyAlaGluLeuAla 155
DB 3 AGAGAATTTTGTGGGCTTCCAAAGAGCAATCAACGAGGTGCTGAAATCTTGCT 62
QY 156 AspThrPheLysAspTyrAlaValSerThrValProValAlaAspGlyLeuHisLeuLys 175
DB 63 GATACCTTTTAAAGACTATGCACTCCACAGTCCCGAGGCGAGATGGGTTCGATTTGAG 122
QY 176 SerPheCysSerMetAlaGlyProAsnLeuLeuAlaIleGlySerSerGluSerAlaGln 195
DB 123 AGTTCTGCAGCATGGCTGGGCTTAACCTGATCGCAATTGGTCTAGTGAATCTGCACAG 182
QY 196 LysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyrAspLysLeuThrValPro 215
DB 183 AAGGGCCCTTAAATATGATCAACAGATGATGACACCGCTATGACAACTCAGTGGCT 242
QY 216 AspAspIle-AlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLe 235
DB 243 GATGACGT-AGCAGCAACTGTATATATCTATATATATATATATATATATATATATAT 301
QY 235 uHisArgThrProGluGluTyrProGluSerAlaLysValTyrGluLysLeuLysAspHi 255
DB 302 GCACCGAACCCGGAAGAGTATCCAGAAGAGTCCAAAGGTTTATGAGAACTGAAGGACCA 361
QY 255 sMetLeuIleProValSerMetSerGluLeuGluLysValAspGlyLeuLeuThrCysCY 275
DB 362 TAGCTGATCCCGGAGGATGCTTGAATGGAAGAGGTGGATGGGCTGCTCAGCTGCTG 421
QY 275 sSerValLeuIleAsnLysLysValAspSer 285

Db 422 CTCAGTTTATTAACAGAAAGTAGACTCC 452
Search completed: June 8, 2004, 02:57:28
Job time : 2570 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 8, 2004, 01:20:43 ; Search time 94 Seconds
(without alignments)
1682.564 Million cell updates/sec

Title: US-09-889-733B-2

Perfect score: 285

Sequence: 1 MAGLGHPSAFGRATHAVRA.....EKVDGLLTCCSVLKNKYDS 285

Scoring table:

OLIGO Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1360442

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xl
-Q/cgn2_1/USPTO spoil_p/US09889733/runat 07062004 083558 8387/app_query.fasta_1.455
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=olip2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09889733 @CGN_1_1_105 @runat 07062004 083558 8387 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=6 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/FCUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	5.3	1397	4	US-09-023-655-37
2	11	3.9	239	4	US-09-702-705-404
3	11	3.9	239	4	US-09-736-457-404
4	11	3.9	239	4	US-09-614-124B-404
5	11	3.9	239	4	US-09-671-325-404
6	11	3.9	239	4	US-09-589-184-404
7	9	3.2	664	1	US-08-455-755-1
8	9	3.2	664	2	US-07-910-260-1
9	9	3.2	664	5	PCT-US91-00878-1
10	9	3.2	774	4	US-09-724-623-43
11	9	3.2	1089	4	US-08-134-000C-1376
12	9	3.2	1212	4	US-09-107-532A-112

13	9	3.2	1596	4	US-09-134-001C-737
14	9	3.2	4253	4	US-08-956-171E-118
15	9	3.2	7360	1	US-08-286-740-1
16	9	3.2	7360	5	PCT-US95-09576-1
17	9	3.2	10011	4	US-08-961-527-76
18	8	2.8	363	6	5182210-19
19	8	2.8	477	4	US-09-489-039A-4729
20	8	2.8	509	6	5182210-15
21	8	2.8	516	4	US-09-252-991A-14927
22	8	2.8	537	4	US-09-252-991A-2859
23	8	2.8	539	6	5182210-13
24	8	2.8	591	6	5182210-17
25	8	2.8	729	6	5182210-21
26	8	2.8	1335	4	US-09-252-991A-14446
27	8	2.8	1386	4	US-09-489-039A-4802
28	8	2.8	1476	4	US-09-252-991A-2659
29	8	2.8	1488	4	US-09-252-991A-2759
30	8	2.8	1522	4	US-09-443-067-25
31	8	2.8	1593	4	US-09-252-991A-5394
32	8	2.8	2181	4	US-09-443-067-19
33	8	2.8	2191	4	US-09-127-219B-1
34	8	2.8	2298	4	US-09-252-991A-3207
35	8	2.8	5190	4	US-09-134-000C-2270
36	8	2.8	4403765	3	US-09-103-840A-2
37	8	2.8	4411529	3	US-09-103-840A-1
38	7	2.5	48	1	US-07-885-689A-8
39	7	2.5	210	4	US-09-540-236-998
40	7	2.5	216	3	US-08-795-430-2
41	7	2.5	216	4	US-09-355-700-2
42	7	2.5	216	4	US-08-601-132-38
43	7	2.5	216	4	US-08-671-573B-38
44	7	2.5	241	4	US-09-857-063-24
45	7	2.5	244	4	US-09-857-063-14

ALIGNMENTS

RESULT 1

US-09-023-655-37

; Sequence 37, Application US/09023655

; Patent No. 6607879

; GENERAL INFORMATION:

; APPLICANT: Cocks, Benjamin G.

; APPLICANT: Susan G. Stuart

; APPLICANT: Jeffrey J. Seilhamer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE

; TITLE OF INVENTION: EXPRESSION

; NUMBER OF SEQUENCES: 1508

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS

; CURRENT APPLICATION DATA: US/09/023,655

; APPLICATION NUMBER: US/09/023,655

; FILING DATE: HEREWITH

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Zeller, Karen J.

; REGISTRATION NUMBER: 37,071

; REFERENCE/DOCKET NUMBER: PA-0001 US

; TELECOMMUNICATION INFORMATION:

```
; TELEPHONE: (650) 855-0855
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1397 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HMCINOT01
; CLONE: 003437
US-09-023-655-37

Alignment Scores:
Pred. No.: 8.49e-07 Length: 1397
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.26% Indels: 0
DB: 4 Gaps: 0

US-09-889-733B-2 (1-285) x US-09-023-655-37 (1-1397)

QY 130 AspValLeuPheThrGlyArgGluPheValGlyLeuSerLys 144
Db 716 GACGTTCTTCACCGCGGAGTTTTCGTAGGCTCTCCAAA 760

RESULT 2
US-09-702-705-404/c
; Sequence 404, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09702,705
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-404

Alignment Scores:
Pred. No.: 0.00359 Length: 239
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.86% Indels: 0
DB: 4 Gaps: 0

US-09-889-733B-2 (1-285) x US-09-702-705-404 (1-239)

QY 130 AspValLeuPheThrGlyArgGluPheVal 140
Db 35 GACGTTCTTCACCGCGGAGTTTTCGTGA 3

RESULT 3
US-09-736-457-404/c
; Sequence 404, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09736,457
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-404

Alignment Scores:
Pred. No.: 0.00359 Length: 239
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.86% Indels: 0
DB: 4 Gaps: 0

US-09-889-733B-2 (1-285) x US-09-736-457-404 (1-239)

QY 130 AspValLeuPheThrGlyArgGluPheVal 140
Db 35 GACGTTCTTCACCGCGGAGTTTTCGTGA 3

RESULT 4
US-09-614-124B-404/c
; Sequence 404, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-404

Alignment Scores:
Pred. No.: 0.00359 Length: 239
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.86% Indels: 0
DB: 4 Gaps: 0

US-09-889-733B-2 (1-285) x US-09-614-124B-404 (1-239)
```


QY 130 AspValleuPheThrGlyArgGluPheVal 140
Db 35 GACGTTCTCTTCACCGCGCGGAGTTTTCGTA 3

RESULT 5
US-09-671-325-404/c
; Sequence 404, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Far, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien

US-09-671-325-404

Alignment Scores:
Pred. No.: 0.00359 Length: 239
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.86% Indels: 0
DB: 4 Gaps: 0

US-09-889-733b-2 (1-285) x US-09-671-325-404 (1-239)

QY 130 AspValleuPheThrGlyArgGluPheVal 140
Db 35 GACGTTCTCTTCACCGCGCGGAGTTTTCGTA 3

RESULT 6
US-09-589-184-404/c
; Sequence 404, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; CURRENT FILING DATE: 2000-06-05
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 404
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapien

US-09-589-184-404

Alignment Scores:
Pred. No.: 0.00359 Length: 239

Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.86% Indels: 0
DB: 4 Gaps: 0

US-09-889-733b-2 (1-285) x US-09-589-184-404 (1-239)

QY 130 AspValleuPheThrGlyArgGluPheVal 140
Db 35 GACGTTCTCTTCACCGCGCGGAGTTTTCGTA 3

RESULT 7
US-08-455-755-1
; Sequence 1, Application US/08455755
; Patent No. 5736345
; GENERAL INFORMATION:
; APPLICANT: Mather, Jennie P.
; APPLICANT: Roberts, Penelope E.
; TITLE OF INVENTION: LUNG CELL LINE AND METHODS OF USE
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Pointe San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,755
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/910260
; FILING DATE: 16-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/00878
; FILING DATE: 08-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/479130
; FILING DATE: 09-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 585PID1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 664 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-455-755-1

Alignment Scores:
Pred. No.: 1.66 Length: 664
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.16% Indels: 0
DB: 1 Gaps: 0

US-09-889-733b-2 (1-285) x US-08-455-755-1 (1-664)

QY 54 LeuGlySerLysLeuGlyLeuGlnVal 62

Db 609 CTCGGTTCTAAGCTTGGGCTGCAGGTC 635
RESULT 8
US-07-910-260-1
; Sequence 1, Application US/07910260
; Patent No. 5830685
; GENERAL INFORMATION:
; APPLICANT: Mather, Jennie P.
; APPLICANT: Roberts, Penelope E.
; TITLE OF INVENTION: Lung Cell Line and Methods of Use
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07910,260
; FILING DATE: 16-JUL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/00878
; FILING DATE: 08-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/479130
; FILING DATE: 9-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 585P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; LENGTH: 664 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-910-260-1
Alignment Scores:
Pred. No.: 1.66 Length: 664
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.16% Indels: 0
DB: 2 Gaps: 0
US-09-889-733B-2 (1-285) x US-07-910-260-1 (1-664)
QY 54 LeuGlySerLysLeuGlyLeuGlnVal 62
Db 609 CTCGGTTCTAAGCTTGGGCTGCAGGTC 635
RESULT 9
PCT-US91-00878-1
; Sequence 1, Application PC/TUS9100878
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Roberts, Penelope E.
; TITLE OF INVENTION: Lung Cell Line And Methods of Use
; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Boulevard
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/00878
; FILING DATE: 19910208
; CLASSIFICATION: 435/240
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S.S.N. 07/479,130
; FILING DATE: 9-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Robert H.
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: 585P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1489
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 664 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US91-00878-1
Alignment Scores:
Pred. No.: 1.66 Length: 664
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.16% Indels: 0
DB: 5 Gaps: 0
US-09-889-733B-2 (1-285) x PCT-US91-00878-1 (1-664)
QY 54 LeuGlySerLysLeuGlyLeuGlnVal 62
Db 609 CTCGGTTCTAAGCTTGGGCTGCAGGTC 635
RESULT 10
US-09-724-623-43
; Sequence 43, Application US/09724623
; Patent No. 6476209
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Lubbers, Mark W
; APPLICANT: Dekker, James
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; TITLE OF INVENTION: them, and methods for using them.
; FILE REFERENCE: 1048U1
; CURRENT APPLICATION NUMBER: US/09/724,623
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Lactobacillus rhamnosus
US-09-724-623-43
Alignment Scores:
Pred. No.: 1.94 Length: 774
Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.16% Indels: 0
DB: 4 Gaps: 0

US-09-889-733B-2 (1-285) x US-09-724-623-43 (1-774)

QY 107 GluAlaLeuGluLysLeuGlnLeuAsn 115
Db 574 GAGGCATTGGAAGTTGCAACTCAAC 600

RESULT 11

US-09-134-000C-1376
; Sequence 1376, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1376
; LENGTH: 1089
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-1376

Alignment Scores:
Pred. No.: 2.76 Length: 1089
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.16% Indels: 0
DB: 4 Gaps: 0

US-09-889-733B-2 (1-285) x US-09-134-000C-1376 (1-1089)

QY 107 GluAlaLeuGluLysLeuGlnLeuAsn 115
Db 988 GAGGCATTGGAAGTTGCAACTTAAT 1014

RESULT 12

US-09-107-532A-112
; Sequence 112, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 112:
SEQUENCE CHARACTERISTICS:
LENGTH: 1212 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...1212
SEQUENCE DESCRIPTION: SEQ ID NO: 112:
US-09-107-532A-112

Alignment Scores:
Pred. No.: 3.09 Length: 1212
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.16% Indels: 0
DB: 4 Gaps: 0

US-09-889-733B-2 (1-285) x US-09-107-532A-112 (1-1212)

QY 107 GluAlaLeuGluLysLeuGlnLeuAsn 115
Db 310 GAGCCTTTAGAAAACTGCAATTGAAT 336

RESULT 13

US-09-134-001C-737
; Sequence 737, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 737
; LENGTH: 1596
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-737

Alignment Scores:
Pred. No.: 4.1 Length: 1596
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.16% Indels: 0
DB: 4 Gaps: 0

US-09-889-733B-2 (1-285) x US-09-134-001C-737 (1-1596)

QY 107 GluAlaLeuGluLysLeuGlnLeuAsn 115

Db 1063 GAAGCTTTAGAAAATTACAACTTAAT 1089
RESULT 14
US-08-956-171E-118
; Sequence 118, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunesch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 118:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4253 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 118:
US-08-956-171E-118
Alignment Scores:
Pred. No.: 11.2 Length: 4253
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.16% Indels: 0
DB: 4 Gaps: 0
US-09-889-733B-2 (1-285) x US-08-956-171E-118 (1-4253)
QY 107 GluAlaLeuGlnLysLeuGlnLeuAsn 115
Db 1148 GAAGCATTAGAAAATTACAACTTAAT 1174
RESULT 15
US-08-286-740-1
; Sequence 1, Application US/08286740
; Patent No. 5561053
; GENERAL INFORMATION:
; APPLICANT: Crowley, Craig W.

; TITLE OF INVENTION: METHOD FOR SELECTING HIGH-EXPRESSING
; TITLE OF INVENTION: HOST CELLS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,740
; FILING DATE: 05-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7360 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-286-740-1
Alignment Scores:
Pred. No.: 19.8 Length: 7360
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.16% Indels: 0
DB: 1 Gaps: 0
US-09-889-733B-2 (1-285) x US-08-286-740-1 (1-7360)
QY 54 LeuGlySerLysLeuGlyLeuGlnVal 62
Db 1563 CTCGGTTCCTAAGCTTGGGCTGCAGGTC 1589
Search completed: June 8, 2004, 02:59:34
Job time : 111 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 8, 2004, 02:18:48 ; Search time 389 Seconds
(without alignments)
3342.336 Million cell updates/sec

Title: US-09-889-733B-2

Perfect score: 285
Sequence: 1 MAGLHPSAFGRATHAVRA.....EKVDGLLTCCSVLNNKVD 285

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2995936 seqs, 2280998010 residues

Word size: 1

Total number of hits satisfying chosen parameters: 5983782

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODE=frame+pz2n.model -DEV=xlp
-Q=/cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:
-DB=Published Applications NA -QFMT=fastap -SUFFIX=olip2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US0989733 -CGN 1 1 723 -runat 07062004 093558_8409
-NCPU=6 -ICPU=3 -NO_WMAP -LARGESQRY -NEG_SCORES=0 -WAIT -DSFBOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

RESULT 1

US-10-450-826-9
; Sequence 9, Application US/10450826
; Publication No. US20040101818A1

GENERAL INFORMATION:

APPLICANT: Ji, Darren
APPLICANT: Axelrod, Douglas W.
APPLICANT: Cook, Jonathan S.
APPLICANT: Jaiswal, Neelam
APPLICANT: Eistein, Richard
APPLICANT: Houghton, Adam
APPLICANT: Mertz, Lawrence
TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiation
FILE REFERENCE: 044921-5039-WO
CURRENT APPLICATION NUMBER: US/10/450,826
CURRENT FILING DATE: 2003-06-18
PRIOR APPLICATION NUMBER: US 60/255,882
PRIOR FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 60/285,691
NUMBER OF SEQ ID NOS: 149
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9

ALIGNMENTS

1	259	90.9	1633	17	US-10-450-826-9	Sequence 9, Appl	
2	259	90.9	4208	17	US-10-450-826-9	Sequence 22, Appl	
3	225	78.9	3778	15	US-10-252-157-328	Sequence 328, App	
4	203	71.2	3651	9	US-09-925-300-512	Sequence 512, App	
5	106	37.2	3008	15	US-10-205-219-83	Sequence 83, Appl	
C	49	17.2	83946	17	US-10-450-826-8	Sequence 8, Appl	
C	39	13.7	3189	10	US-09-814-353-20553	Sequence 20553, A	
8	20	7.0	60	10	US-09-908-975-9030	Sequence 9030, A	
9	15	5.3	452	10	US-09-918-995-14871	Sequence 14871, A	
10	15	5.3	472	10	US-09-918-995-9117	Sequence 9117, Ap	
11	15	5.3	1397	17	US-10-641-643-37	Sequence 37, Appl	
C	12	3.9	239	9	US-09-736-457-404	Sequence 404, App	
C	13	3.9	239	9	US-09-902-941-404	Sequence 404, App	
C	14	3.9	239	9	US-09-849-626-404	Sequence 404, App	
C	15	3.9	239	10	US-09-476-300-404	Sequence 404, App	
C	16	3.9	239	13	US-10-283-017-404	Sequence 404, App	
C	17	3.9	239	15	US-10-017-754-404	Sequence 404, App	
C	18	3.9	239	15	US-10-113-872-404	Sequence 404, App	
19	10	3.5	332	9	US-09-783-590-5309	Sequence 5309, Ap	
20	10	3.5	770	9	US-09-910-943-9	Sequence 9, Appl	
21	9	3.2	456	9	US-09-815-242-9440	Sequence 9440, Ap	
22	9	3.2	640	15	US-10-005-675-18	Sequence 18, Appl	
C	23	3.2	645	13	US-10-027-632-215186	Sequence 215186	
C	24	3.2	645	16	US-10-027-632-215186	Sequence 215186,	
25	9	3.2	664	15	US-10-005-675-16	Sequence 16, Appl	
26	9	3.2	664	15	US-10-005-675-17	Sequence 17, Appl	
27	9	3.2	774	15	US-10-288-930-43	Sequence 43, Appl	
C	28	9	3.2	1035	13	US-10-282-122A-11501	Sequence 11501, A
C	29	9	3.2	4253	8	US-08-781-986A-118	Sequence 118, App
30	9	3.2	4253	13	US-10-329-624-118	Sequence 118, App	
31	9	3.2	6491	9	US-09-070-927A-156	Sequence 156, App	
C	32	9	10011	13	US-10-158-844-76	Sequence 76, Appl	
33	9	3.2	69168	13	US-10-087-192-124	Sequence 124, App	
34	8	2.8	60	10	US-09-908-975-9031	Sequence 9031, Ap	
35	8	2.8	290	13	US-10-424-599-86153	Sequence 86153, A	
36	8	2.8	403	10	US-09-918-995-19254	Sequence 19254, A	
37	8	2.8	445	10	US-09-918-995-5286	Sequence 5286, Ap	
C	38	2.8	453	13	US-10-424-599-21791	Sequence 21791, A	
C	39	2.8	471	13	US-10-424-599-120762	Sequence 120762,	
40	8	2.8	500	11	US-09-864-408A-8175	Sequence 8175, Ap	
41	8	2.8	588	13	US-10-027-632-281162	Sequence 281162,	
42	8	2.8	588	16	US-10-027-632-281162	Sequence 281162,	
C	43	2.8	628	13	US-10-027-632-12419	Sequence 12419, A	
C	44	2.8	628	16	US-10-027-632-12419	Sequence 12419, A	
C	45	2.8	632	13	US-10-027-632-104847	Sequence 104847,	

```

; LENGTH: 1633
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. AB001915
; NAME/KEY: unsure
; LOCATION: (1)..(1633)
; OTHER INFORMATION: n = a or c or g or t
US-10-450-826-9

Alignment Scores:
Pred. No.:      8,086-264      Length:      1633
Score:          259.00        Matches:      259
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:      90.88%      Indels:        0
DB:              17          Gaps:          0

US-09-889-733B-2 (1-285) x US-10-450-826-9 (1-1633)
QY 27 GlnHisAlaLeuArgSerAlaLysGlyGluValAspValAlaArgAlaGluArgGln 46
DB 401 CAGCAGCGCTGAGAAAGCCGAGGCGGAGGAGGTGGACGTGCGCCGCGCGGACGCGAG 460
QY 47 HisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuPro 66
DB 461 CACCAGCTCTACGTGGCGGTGCTGGGCGACGAACTGGGGCTGCGAGGTGGTGGAGCTGCCG 520
QY 67 AlaAspGluSerLeuProAspCysValPheValGluAspValAlaValAlaValCysGluGlu 86
DB 521 GCCGACGAGAGCTTCGCGACGTGCTCTTCGTGGAGGACGTGGCGCTGGTGGGAGGAG 580
QY 87 ThrAlaLeuLeuThrArgProGlyAlaProSerArgArgLysGlyValAspMetMetLys 106
DB 581 ACCGGCCCTCATCACCAGCCCGGGCGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 640
QY 107 GluAlaLeuGluLysLeuGlnLeuAsnLeuValGluMetLysAspGluAsnAlaThrLeu 126
DB 641 GAAGCATTTAGAAAACCTTCAGCTCAATATAGTAGAGATGAAAGATGAAATGCAACTTTA 700
QY 127 AspGlyGlyAspValLeuPheThrGlyArgGluPheValGlyLeuSerLysArgThr 146
DB 701 GATGCGGAGAGTGTATTATTCACAGCGCAGAGAAATTTTGTGGGCTTTTCCAAAGGACA 760
QY 147 AsnGlnArgGlyAlaGluLeuLeuAlaAspThrPheLysAspTyrAlaValSerThrVal 166
DB 761 AATCAACGAGGTGCTGAATCTTGGCTGATCTTTTAAGGACTATGCACTCTCCACAGTG 820
QY 167 ProValAlaAspGlyLeuHisLeuLysSerPheCysSerMetAlaGlyProAsnLeuLeu 186
DB 821 CCAGTGGCAGATGGGTGCAATTTGAAGAGTTCTGCAGCATGGCTGGGCGCTTAACCTGATC 880
QY 187 AlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAsp 206
DB 881 GCAATTGGGTCTAGTGAATCTGCACAGAGGCGCTTAAGATCATGCAACAGATGAGTGAC 940
QY 207 HisArgTyrAspLysLeuThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsn 226
DB 941 CACCGCTACGAAACTCACTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1000
QY 227 IleProAsnLysGlyHisValLeuLeuHisArgThrProGluGluTyrProGluSerAla 246
DB 1001 ATCCCCAACAAAGGCGACGTCTTGTGACCGAACCCCGGAGAGATATCCAGAAAGTGCA 1060
QY 247 LysValTyrGluLysLeuLysAspHisMetLeuLeuProValSerMetSerGluLeuGlu 266
DB 1061 AAGGTTTATGAGAAACTGAGGAGGACCATATGCTGATCCCCCGTGGATGATGCTGGAAGTGA 1120
QY 267 LysValAspGlyLeuLeuThrCysCysSerValLeuLeuAsnLysLysValAspSer 285
DB 1121 AAGGTGATGGGCTGCTCACTGCTGCTCAGTTTATTAACAGAAAGTAGACTCC 1177
```

```

RESULT 2
US-10-468-125-22
; Sequence 22, Application US/10468125
; Publication No. US20040082061A1
; GENERAL INFORMATION:
; APPLICANT: ASTROMOFF, Anna
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: DING, Li
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: FORSYTHE, Ian J.
; APPLICANT: GIETZEN, Kimberly J.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: LU, Yan
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: RING, Huijun Z.
; APPLICANT: SANJANWALA, Madhusudan
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: WARREN, Bridget A.
; APPLICANT: YUE, Yuming
; APPLICANT: YUE, Henry
; APPLICANT: ZEBARJADIAN, Yeganeh
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: FI-0363 USN
; CURRENT APPLICATION NUMBER: US/10/468,125
; CURRENT FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: PCT/US02/04918
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,643
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/271,332
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/276,767
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: US 60/282,077
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/285,447
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/287,060
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/288,543
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 22
; LENGTH: 4208
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7671089CBI
US-10-468-125-22
```

```

Alignment Scores:
Pred. No.:      1,896-263      Length:      4208
Score:          259.00        Matches:      259
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:      90.88%      Indels:        0
DB:              17          Gaps:          0

US-09-889-733B-2 (1-285) x US-10-468-125-22 (1-4208)
QY 27 GlnHisAlaLeuArgSerAlaLysGlyGluValAspValAlaArgAlaGluArgGln 46
DB 446 CACACCGCTGAGAAAGCCGAGGCGGAGGAGGTGGACGTGCGCCGCGCGGACGCGAG 505
QY 47 HisGlnLeuTyrValGlyValLeuGlySerLysLeuGlyLeuGlnValValGluLeuPro 66
DB 506 CACCAGCTCTACGTGGCGGTGCTGGGCGACGAACTGGGGCTGCGAGGTGGTGGAGCTGCCG 565
```



```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 512
; LENGTH: 3651
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1283)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3641)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3650)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-512

Alignment Scores:
Pred. No.: 2,85e-204 Length: 3651
Score: 203.00 Matches: 203
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 71.23% Indels: 0
DB: 9 Gaps: 0

US-09-889-733B-2 (1-285) x US-09-925-300-512 (1-3651)

QY 73 AspCysValPheValGluAspValAlaValValCysGluGluThrAlaLeuLeuThrArg 92
DB 4 GACTCGCTTCGTGGAGAGCGTGGCTGGTGGAGGAGCGGCGCTCATCACCGA 63
QY 93 ProGlyAlaProSerArgArgLysGluValAspMetMetLysGluAlaLeuGluLysLeu 112
DB 64 CCGCGGCGCGAGCGGAGGAGGAGGTTGACATGATGAAAGAGCATTAAGAAAACTT 123
QY 113 GlnLeuAsnIleValGluMetLysAspGluAsnAlaThrLeuAspGlyGlyAspValLeu 132
DB 124 CAGCTCAATATAGTAGATGAAGATGAATGAATGCACTTTAGATGGCGGAGATGTTTA 183
QY 133 PheThrGlyArgGluPhePheValGlyLeuSerLysArgThrAsnGlnArgGlyValGlu 152
DB 184 TTCACAGCAGAGATTTTGTGGGCTTTCCAAAGGACCAATCAACGAGGTGCTGAA 243
QY 153 IleLeuAlaAspThrPheLysAspTyrAlaValSerThrValProValAlaAspGlyLeu 172
DB 244 ATCTTGGCTGATACATTTAAAGACTATGCTGCTCCACAGTCCGAGTGGCAGATGGGTTG 303
QY 173 HisLeuLysSerPheCysSerMetAlaGlyProAsnLeuIleAlaIleGlySerSerGlu 192
DB 304 CATTTGAAGAGTTTCTGCAGCATGCTGGGCTTAACCTGATCGCAATTTGGTCTAGTGAA 363
QY 193 SerAlaGlnLysAlaLeuLysIleMetGlnGlnMetSerAspHisArgTyrAspLysLeu 212
DB 364 TCTGCACAGAGCGCTTAAGATCATGCAACAGATGAGTGACCAACCGCTACGACAAACTC 423
QY 213 ThrValProAspAspIleAlaAlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHis 232
DB 424 ACTGTCCTGATGATAGCAGCAAACTGTATATCTAAATATCCCCCAAAAGGGGAC 483
QY 233 ValLeuLeuHisArgThrProGluTyrProGluSerAlaLysValTyrGluLysLeu 252
DB 484 GTCCTGCTGCACCAACCGGAGAGATCCAGAAAGTGCAAAAGTTTATGAGAAACTG 543
QY 253 LysAspHisMetLeuIleProValSerMetSerGluLeuGluLysValAspGlyLeuLeu 272
DB 544 AAGGACCATATGCTGATCCCCGTGAGCATGTGTAACCTGGAAGAGGTGGTGGGCTGCTC 603
QY 273 ThrCysCys 275
DB 604 ACTGCTGT 612

RESULT 5
US-10-205-219-83

; Sequence 83, Application US/10205219
; Publication No. US20030138803A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pincock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018200
; CURRENT APPLICATION NUMBER: US/10/205,219
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 83
; LENGTH: 3008
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: N-G, N-G dimethylarginine dimethylaminohydrolase
; NAME/KEY: misc_feature
; LOCATION: 2411, 2412
; OTHER INFORMATION: n is a or g or c or t
US-10-205-219-83

Alignment Scores:
Pred. No.: 9,57e-102 Length: 3008
Score: 106.00 Matches: 152
Percent Similarity: 98.70% Conservative: 0
Best Local Similarity: 98.70% Mismatches: 1
Query Match: 37.19% Indels: 2
DB: 15 Gaps: 0

US-09-889-733B-2 (1-285) x US-10-205-219-83 (1-3008)

QY 65 LeuProAlaAspLysSerLeuProAspCysValPheValGluAspValAlaValValCys 84
DB 624 CTGCGCGCGCAGCAGAGCGCTGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 683
QY 85 GluGluThrAlaLeuIleThrArgProGlyAlaProSerArgArgLysGluValAspMet 104
DB 684 GAGGAGACCGCTTCATACCCCGCGCGCGCGCTAGCCGCGAGAGAGGTTGACATG 743
QY 105 MetLysGluAlaLeuGluLysLeuGlnLeuAsnIleValGluMetLysAspGluAsnAla 124
DB 744 ATCAAGAGGCTTTGGAAAACTTCAGCTCAACATAGTAGATGAAGATGAAGATGA 803
QY 125 ThrLeuAspGlyGlyAspValLeuPheThrGlyArgGluPhePheValGlyLeuSerLys 144
DB 804 ACCTTAGATGGTGGGACGCTCTATTTCACAGCAGAGAGTCTTTTGTGGGCTTTTCCAAA 863
QY 145 ArgThrAsnGlnArgGlyAlaGluIleLeuAlaAspThrPheLysAspTyrAlaValSer 164
DB 864 AGGACAAATCAACGAGGTGCTGAGATCTTTGGCTGATACCTTCAAGGACTACGAGTTTCC 923
QY 165 ThrValProValAlaAspGly-LeuHisLeuLysSerPheCysSerMetAlaGlyProAs 184
DB 924 ACAGTCCCGCTGGCCGATTC-TTTGCAATTTAAAGAGTCTCTGAGCATGGCTGGGCCCAA 982
QY 184 nLeuIleAlaIleGlySerSerGluSerAlaGlnLysAlaLeuLysIleMetGlnGlnMe 204
DB 983 CTTGATCGCAATAGGTCCTGAGTGAATCTGCGGAGAGGCGCTCAAGATCATGCAACAGAT 1042
QY 204 tSerAspHisArgTyrAspLysLeuThrValProAspAsp 217
DB 1043 GAGTGACCGCTTATGACAAAGCTCACTGTATCCGGACGAC 1082

RESULT 6
US-10-450-826-8/c
; Sequence 8, Application US/10450826
```



```
; Publication No. US2004010181A1
; GENERAL INFORMATION:
; APPLICANT: JI, Darren
; APPLICANT: Axelrod, Douglas W.
; APPLICANT: Cook, Jonathan S.
; APPLICANT: Jaiswal, Neelam
; APPLICANT: Bistein, Richard
; APPLICANT: Houghton, Adam
; APPLICANT: Mertz, Lawrence
; TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiation
; FILE REFERENCE: 044921-5039-WO
; CURRENT APPLICATION NUMBER: US/10/450,826
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 60/255,882
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/285,691
; PRIOR FILING DATE: 2001-04-24
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 8
; LENGTH: 83946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. AL078459
; US-10-450-826-8

Alignment Scores:
Pred. No.: 3,686-40 Length: 83946
Score: 49.00 Matches: 49
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 17.19% Indels: 0
DB: 17 Gaps: 0

US-09-889-733B-2 (1-285) x US-10-450-826-8 (1-83946)

QY 200 IleMerGlnGluMetSerAspHisArgTyrAspLysLeuThrValProAspAspIleAla 219
Db 18610 ATCATGCAACAGATGAGTACCCACCCCTACGAAACTCCTGCTGCTGATGACATAGCA 18551

QY 220 AlaAsnCysIleTyrLeuAsnIleProAsnLysGlyHisValLeuLeuHisArgThrPro 239
Db 18550 GCAAACTGTATATATCTAAATATCCCAACAAAGGCGACGCTTGTGTCACCGACCCCG 18491

QY 240 GluGluTyrProGluSerAlaLysVal 248
Db 18490 GAAGAGTATCCAGAAAGTGCAGGTA 18464

RESULT 7
US-09-814-353-20553/c
; Sequence 20553, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25

; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20553
; LENGTH: 3189
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3189
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-20553

Alignment Scores:
Pred. No.: 7,46e-31 Length: 3189
Score: 39.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.68% Indels: 0
DB: 10 Gaps: 0

US-09-889-733B-2 (1-285) x US-09-814-353-20553 (1-3189)

QY 247 LysValTyrGluLysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGlu 266
Db 3186 AAGTTTATGAGAAACTGAAGACCATATGCTGATCCCTGACCATGCTGAACTGGAA 3127

QY 267 LysValAspGlyLeuLeuThrCysCysSerValLeuIleAsnLysLysValAspSer 285
Db 3126 AAGGTGGATGGCTGCTCACCTGCTGCTCAGTTTATTAACAAGAAAGTAGACTCC 3070

RESULT 8
US-09-908-975-9030
; Sequence 9030, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 9030
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-908-975-9030

Alignment Scores:
Pred. No.: 2,67e-12 Length: 60
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.02% Indels: 0
DB: 10 Gaps: 0

US-09-889-733B-2 (1-285) x US-09-908-975-9030 (1-60)

QY 251 LysLeuLysAspHisMetLeuIleProValSerMetSerGluLeuGluLysValAspGly 270
Db 1 AAAGTGAAGGACCATATGCTGATCCCGTGAGCATGCTGAACTGGAAGAGGTGGATGGG 60
```

RESULT 9

US-09-918-995-14871
; Sequence 917, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14871
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(452)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-14871

Alignment Scores:
Pred. No.: 3,16e-06 Length: 452
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.26% Indels: 0
Gaps: 10

US-09-889-733B-2 (1-285) x US-09-918-995-14871 (1-452)

Qy 130 AspValleupheThrGlyArgGluPheValGlyLeuSerIys 144
Db 324 GACGTTCTCTTCACGGCGGGAGTTTTTCGTAGGCTCTCCAAA 368

RESULT 10

US-09-918-995-9117
; Sequence 917, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9117
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-9117

Alignment Scores:
Pred. No.: 3,29e-06 Length: 472
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.26% Indels: 0
Gaps: 10

US-09-889-733B-2 (1-285) x US-09-918-995-9117 (1-472)

Qy 130 AspValleupheThrGlyArgGluPheValGlyLeuSerIys 144
Db 357 GACGTTCTCTTCACGGCGGGAGTTTTTCGTAGGCTCTCCAAA 401

RESULT 11

US-10-641-643-37
; Sequence 37, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; Susan G. Stuart
; Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/641,643
; FILING DATE: 14-Aug-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1397 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HMCINOT01
; CLONE: 003437
; SEQUENCE DESCRIPTION: SEQ ID NO: 37 :

US-10-641-643-37
Pred. No.: 8,68e-06 Length: 1397
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.26% Indels: 0
Gaps: 17

US-09-889-733B-2 (1-285) x US-10-641-643-37 (1-1397)

Qy 130 AspValleupheThrGlyArgGluPheValGlyLeuSerIys 144
Db 716 GACGTTCTCTTCACGGCGGGAGTTTTTCGTAGGCTCTCCAAA 760

RESULT 12

US-09-736-457-404/C
; Sequence 404, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:

APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedwick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 404
LENGTH: 239
TYPE: DNA
ORGANISM: Homo sapien
US-09-736-457-404

Alignment Scores:
Pred. No.: 239
Score: 11.00
Length: 239
Matches: 11
Percent Similarity: 100.00%
Conservative: 0
Best Local Similarity: 100.00%
Mismatch: 0
Indels: 0
Query Match: 3.86%
Gaps: 0
DB:

US-09-889-733B-2 (1-285) x US-09-736-457-404 (1-239)

QY 130 AspValLeuPheThrGlyArgGluPheVal 140
DB 35 GACGTTCTCTTCCACCGCGGGAGTTTTCGTA 3

RESULT 13

US-09-902-941-404/c
Sequence 404, Application US/09902941
Patent No. US2002017952A1
GENERAL INFORMATION:
APPLICANT: Henderson, Robert A.
APPLICANT: Wang, Tongtong
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Johnson, Jeffrey C.
APPLICANT: Retter, Marc W.
APPLICANT: Marnerakis, Margarita
APPLICANT: Carter, Darick
APPLICANT: Fanger, Gary R.
APPLICANT: Vedwick, Thomas S.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: McNabb, Andria
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C17
CURRENT APPLICATION NUMBER: US/09/902,941
CURRENT FILING DATE: 2001-07-10
NUMBER OF SEQ ID NOS: 2002
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 404
LENGTH: 239
TYPE: DNA
ORGANISM: Homo sapiens
US-09-902-941-404

Alignment Scores:
Pred. No.: 239
Score: 11.00
Length: 239
Matches: 11
Percent Similarity: 100.00%
Conservative: 0
Best Local Similarity: 100.00%
Mismatch: 0
Indels: 0
Query Match: 3.86%
Gaps: 0
DB:

US-09-889-733B-2 (1-285) x US-09-902-941-404 (1-239)
QY 130 AspValLeuPheThrGlyArgGluPheVal 140
DB 35 GACGTTCTCTTCCACCGCGGGAGTTTTCGTA 3

RESULT 14

US-09-849-626-404/c
Sequence 404, Application US/09849626
Publication No. US20020197669A1
GENERAL INFORMATION:
APPLICANT: Bangur, Chaitanya
APPLICANT: Fanger, Gary
APPLICANT: Wang, Aijun
APPLICANT: Wang, Tongtong
APPLICANT: Switzer, Anne
APPLICANT: McNeill, Patricia
APPLICANT: Clapper, Jonathan
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C16
CURRENT APPLICATION NUMBER: US/09/849,626
CURRENT FILING DATE: 2001-05-03
NUMBER OF SEQ ID NOS: 1926
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 404
LENGTH: 239
TYPE: DNA
ORGANISM: Homo sapien
US-09-849-626-404

Alignment Scores:
Pred. No.: 0.0304
Score: 11.00
Length: 239
Matches: 11
Percent Similarity: 100.00%
Conservative: 0
Best Local Similarity: 100.00%
Mismatch: 0
Indels: 0
Query Match: 3.86%
Gaps: 0
DB:

US-09-889-733B-2 (1-285) x US-09-849-626-404 (1-239)

QY 130 AspValLeuPheThrGlyArgGluPheVal 140
DB 35 GACGTTCTCTTCCACCGCGGGAGTTTTCGTA 3

RESULT 15

US-09-476-300-404/c
Sequence 404, Application US/09476300
Publication No. US20030125245A1
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C3
CURRENT APPLICATION NUMBER: US/09/476,300
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 785
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 404
LENGTH: 239
TYPE: DNA
ORGANISM: Homo sapien
US-09-476-300-404

Alignment Scores:
Pred. No.: 0.0304
Score: 11.00
Length: 239
Matches: 11
Percent Similarity: 100.00%
Conservative: 0
Best Local Similarity: 100.00%
Mismatch: 0
Indels: 0
Query Match: 3.86%
Gaps: 0
DB:

US-09-889-733B-2 (1-285) x US-09-476-300-404 (1-239)
Qy 130 AspValLeuPheThrGlyArgGluPhePheVal 140
Db 35 GACGTTCTCTCACCGCGCGGAGTTTTCGTA 3

Search completed: June 8, 2004, 03:06:40
Job time : 423 secs